



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 110595

TO: Karen A Lacourciere
Location: CM1/11D09
Art Unit: 1635
Tuesday, December 23, 2003
Case Serial Number: 09744679

From: Paul Schulwitz
Location: Biotech-Chem Library
CM1-6B06
Phone: 305-1954

paul.schulwitz@uspto.gov

Search Notes

Examiner Lacourciere,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(703)305-1954

STIC-Biotech/ChemLib

110595

RECEIVED

From: Lacourciere, Karen
Sent: Tuesday, December 16, 2003 5:38 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request 09/744,679

DEC 17 2003

(STIC)

Please search SEQ ID NO:7 for 09/744,679

Thank-you!

Karen A. Lacourciere Ph.D.

CM1 11D09 GAU 1635
(703) 308-7523

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 12/23
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2003, 17:03:24 ; Search time 236 Seconds
(without alignments)
6268.187 Million cell updates/sec

Title: US-09-744-679-7
Perfect score: 548
Sequence: 1 cgccgaggtcgttgacc.....ccatcgttgttcacctg 548

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues 5105512
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	548	100.0	548	21	AAZ50772 Human thrombin rec
2	548	100.0	3472	20	AAZ32191 Human thrombin rec
3	548	100.0	3472	25	ACA56730 Human signalling p
4	548	100.0	3480	13	AAQ28568 Human thrombin rec
5	548	100.0	3480	21	AAZ50771 Human thrombin rec
6	548	100.0	3592	25	ACC50141 Breast cancer asso
7	548	100.0	3592	25	ABZ42767 Human thrombin rec
8	387	70.6	1278	24	ABK70889 Human cDNA encodin

9	328.4	59.9	2910	18	AAZ50772	ThR-GPA1 fusion ge
10	322.4	58.8	6546	21	AAF21433	Human low adenosin
11	322.4	58.8	6546	21	AAA35311	Human adenosine re
12	321	58.6	3182	21	AAF21432	Human low adenosin
13	321	58.6	3182	21	AAA35310	Human adenosine re
14	282.4	51.5	1764	15	AAQ73590	Fragment of the hu
15	249	45.4	772	23	AAS68315	DNA encoding novel
16	240.8	43.9	1209	24	ABK70887	Human cDNA encodin
17	184.2	33.6	1312	18	AAZ93367	Mouse thrombin rec
18	148.2	27.0	6436	24	ABL32681	Human immune syste
19	147	26.8	1116	24	ABK70888	Human cDNA encodin
20	146.6	26.8	6436	24	ABL32680	Human immune syste
21	146	26.6	572	21	AAA27044	Human cell surface
22	137.2	25.0	556	24	ABQ23658	Oligonucleotide fo
23	137.2	25.0	556	24	ABQ23659	Oligonucleotide fo
24	135.6	24.7	556	24	ABQ23660	Oligonucleotide fo
25	135.6	24.7	556	24	ABQ23661	Oligonucleotide fo
26	126.6	23.1	1300	19	AAV33372	PROLAC FLAG-conne
27	125.4	22.9	1300	19	AAV33371	PROLAC FLAG-conne
28	56	10.2	56	19	AAV33375	Fragment of thromb
29	55	10.0	55	19	AAV33376	Fragment of thromb
30	48.8	8.9	47981	22	AAF30757	Micromonospora meg
31	46	8.4	54	19	AAV33374	Fragment of thromb
32	44.6	8.1	192	18	AAZ97178	Thrombin site-DR b
33	44.4	8.1	12111	22	ABA21422	Human nervous syst
34	42	7.7	3272	24	ABI99475	Mouse ischaemic co
35	41.8	7.6	2259	23	AAS79424	DNA encoding novel
36	41.2	7.5	1362	25	ABX56041	M. echinospora cal
37	40.8	7.4	67212	21	AAA08954	WFS1 variant genom
38	40.6	7.4	189	18	AAZ97177	Thrombin site-DR a
39	40.6	7.4	2341	19	AAV62155	HSV-2 strain SB5 C
40	40.6	7.4	2694	19	AAV62145	HSV-2 strain SB5 C
41	40.6	7.4	117213	19	AAV62176	HSV-2 strain SB5 C
42	40.6	7.4	154746	24	AAD25519	Human herpesvirus
43	40.4	7.4	327	23	AAH88285	CNS disorder-relat
44	40.2	7.3	32195	22	AAS36105	Human cardiovascular
45	40.2	7.3	32195	22	AAS31532	Human DNA for a no

ALIGNMENTS

RESULT 1	
AAZ50772	
ID	AAZ50772 standard; cDNA; 548 BP.
XX	
AC	AAZ50772;
XX	
DT	31-MAY-2000 (first entry)
XX	
DE	Human thrombin receptor antisense cDNA.
XX	
KW	Human; thrombin receptor; ThR; PAR-1; protease activated receptor;
KW	antisense molecule; PAR antibody; cytostatic; therapeutic;
KW	metastatic tumour cell; placental implantation; invasive cell; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200008150-A1.
XX	
PD	17-FEB-2000.
XX	
PF	05-FEB-1999; 99WO-IL00079.
XX	
PR	07-AUG-1998; 98IL-0125698.
XX	
PA	(HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX	
PI	Bar-Shavit R;
XX	
DR	WPI; 2000-205706/18.
XX	
PT	Treating metastatic tumor cells useful for treating disorders involving

PT placenta implantation in a female comprises administration of an
PT antisense molecule complementary to an RNA sequence of a protease
PT activated receptor protein -
XX
PS Claim 6; Fig 2; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using
CC an antisense molecule comprising a polynucleotide complementary to an
CC RNA sequence of a protease activated receptor (PAR) protein, or an
CC antibody capable of binding to a PAR protein. The antisense molecules and
CC antibodies of PAR protein are also used to treat disorders associated
CC with implantation of placenta. The present sequence is an antisense
CC cDNA of human thrombin receptor (ThR) (also known as PAR-1), which is a
CC G-coupled protein belonging to the PAR family. This sequence is useful
CC in the treatment of invasive cells.

XX SQ Sequence 548 BP; 122 A; 166 C; 148 G; 112 T; 0 other;

Query Match 100.0%; Score 548; DB 21; Length 548;
Best Local Similarity 100.0%; Pred. No. 6.3e-139;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGCCACCCCTGCTCTGCCCTGCCGCG 60
Db |||||
1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGCCACCCCTGCTCTGCCCTGCCGCG 60
QY 61 AAGACCGGCTCCCGACCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
Db |||||
61 AAGACCGGCTCCCGACCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
QY 121 GGGGAGCGCTCCCGGAGCAGCGCCGCGAGAGCGCGGACAAATGGGGCGCGCGGCTGC 180
Db |||||
121 GGGGAGCGCTCCCGGAGCAGCGCCGCGAGAGCGCGGACAAATGGGGCGCGCGGCTGC 180
QY 181 TGCTGGTGGCGCGCTGCTTCACTGCTGTGCGGCGCGCTGTTGTCTGCCGACCCCGGCCC 240
Db |||||
181 TGCTGGTGGCGCGCTGCTTCACTGCTGTGCGGCGCGCTGTTGTCTGCCGACCCCGGCCC 240
QY 241 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGGA 300
Db |||||
241 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGGA 300
QY 301 ACCCAATGATAAATATGACCAATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
Db |||||
301 ACCCAATGATAAATATGACCAATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 420
Db |||||
361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 420
QY 421 TCATCTCAGAAGATGCTCCGATATTTGACCAAGTCTCTGCTGACACTTTTGTCCCAT 480
Db |||||
421 TCATCTCAGAAGATGCTCCGATATTTGACCAAGTCTCTGCTGACACTTTTGTCCCAT 480
QY 481 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db |||||
481 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
QY 541 TCATCTCTG 548
Db |||||
541 TCATCTCTG 548

RESULT 2
AAZ32191
ID PAZ32191 standard; cDNA; 3472 BP.

XX AAZ32191;
AC
XX
DT 13-JAN-2000 (first entry)
XX
DE Human thrombin receptor nucleotide sequence.
XX

KW Human; coding sequence polymorphism; vascular pathology gene;
KW polymorphic site; phenotype correlation; forensic; paternity testing;
KW medicine; genetic analysis; vascular disease; ds.

XX Homo sapiens.

XX WO9950454-A2.

XX 07-OCT-1999.

XX 26-MAR-1999; 99WO-US06473.

XX 01-APR-1998; 98US-0054272.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

XX WPI; 1999-620066/53.

XX P-PSDB; AAY49570.

XX Determination of polymorphisms in genes, especially those identifying
PT predisposition to vascular disease -

XX Claim 1; Fig 35; 134pp; English.

XX AAZ32159 to AAZ32194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AAY49550 to AAY49573 represent the proteins which correspond
CC to some of the reference alleles.

SQ Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;

Query Match 100.0%; Score 548; DB 20; Length 3472;
Best Local Similarity 100.0%; Pred. No. 1.3e-138;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGCCACCCCTGCTCTGCCCTGCCGCG 60
Db |||||
64 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGCCACCCCTGCTCTGCCCTGCCGCG 123
QY 61 AAGACCGGCTCCCGACCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
Db |||||
124 AAGACCGGCTCCCGACCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 183
QY 121 GGGGAGCGCTCCCGGAGCAGCGCCGCGAGAGCGCCGAGACAAATGGGGCGCGCGGCTGC 180
Db |||||
184 GGGGAGCGCTCCCGGAGCAGCGCCGCGAGAGCGCCGAGACAAATGGGGCGCGCGGCTGC 243
QY 181 TGCTGGTGGCGCGCTGCTTCACTGCTGTGCGGCGCGCTGTTGTCTGCCGACCCCGGCCC 240
Db |||||
244 TGCTGGTGGCGCGCTGCTTCACTGCTGTGCGGCGCGCTGTTGTCTGCCGACCCCGGCCC 303
QY 241 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGGA 300
Db |||||
304 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGGA 363
QY 301 ACCCAATGATAAATATGACCAATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
Db |||||
364 ACCCAATGATAAATATGACCAATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 420
Db |||||
424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 483
QY 421 TCATCTCAGAAGATGCTCCGATATTTGACCAAGTCTCTGCTGACACTTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTGACGAGTCTCTGGTGACACTCTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 3

ACA56730

ID ACA56730 standard; cDNA; 3472 BP.

XX

AC

XX

DT 06-JUN-2003 (first entry)

XX Human signalling pathway polynucleotide probe SEQ ID NO 1328.

DE

XX

KW Human; probe; ss; array element; Parkinson's disease;

KW signalling pathway population; cancer; adenocarcinoma; leukaemia;

KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

XX Homo sapiens.

XX US6500938-B1.

PN 31-DEC-2002.

XX 30-JAN-1998; 98US-0016434.

PF 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

PA Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

DR

XX Combination of polynucleotide probes, useful as array elements in a

PT microarray for monitoring the expression of a number of target

PT polynucleotides -

XX Claim 1; SEQ ID NO 1328; 65pp; English.

PS

XX The invention relates to a combination which, comprises a number of

CC polynucleotide probes comprising a sequence selected from one of the 1490

CC sequences mentioned in the specification. The combination is useful as an

CC array element in a microarray for monitoring the expression of a number

CC of target polynucleotides. The microarray is particularly useful in the

CC diagnosis and treatment of cancer and immunopathology and neuropathology.

CC The microarray is useful in diagnostics and treatment regimens, drug

CC discovery and development, toxicological and carcinogenicity studies,

CC forensics and pharmacogenomics. The microarray is also useful for

CC monitoring progression of diseases and for developing sophisticated

CC profiles for the effects of currently available therapeutic drugs. The

CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs

CC and genomic fragments and in research and diagnostic applications. The

CC array can detect changes in expression in a large number of genes coding

CC for different signaling pathway populations which can be used to diagnose

CC various diseases including cancer e.g. adenocarcinoma and leukaemia,

CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

CC and Parkinson's disease. The present sequence represents a polynucleotide

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format directly from USPTO

CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.

XX

SQ Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;

Query Match 100.0%; Score 548; DB 25; Length 3472;

Best Local Similarity 100.0%; Pred. No. 1.3e-138;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTTGGGACCCCTGCGCTCTGCGTCCGCG 60

Db 64 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTTGGGACCCCTGCGCTCTGCGTCCGCG 123

QY 61 AAGACCCGGCTCCCGACCCCGCAGAGTCAGGAGAGAGGGTGAAGCGAGCAGCCCGAGGC 120

Db 124 AAGACCCGGCTCCCGACCCCGCAGAGTCAGGAGAGAGGGTGAAGCGAGCAGCCCGAGGC 183

QY 121 GGGGAGCCTCCCGGAGCAGCGCCCGCAGAGCCCGGACCAATGGGCGCGGCGCTGC 180

Db 184 GGGGAGCCTCCCGGAGCAGCGCCCGCAGAGCCCGGACCAATGGGCGCGGCGCTGC 243

QY 181 TGCTGGTGGCGCTGCTTCACTCTGTGCGGCGCGCTGTTGTCTGCGCGCACCCGGGCCC 240

Db 244 TGCTGGTGGCGCTGCTTCACTCTGTGCGGCGCGCTGTTGTCTGCGCGCACCCGGGCCC 303

QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCCGGTCAATTCTTCTCAGGA 300

Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCCGGTCAATTCTTCTCAGGA 363

QY 301 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGGTAA 360

Db 364 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGGTAA 423

QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTTCTTCAAAAACAACTTCTTCCAT 420

Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTTCTTCAAAAACAACTTCTTCCAT 483

QY 421 TCATCTCAGAAGATGCTCCGGATATTGACCAAGTCTGGCTGACACTCTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTGACCAAGTCTGGCTGACACTCTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 4

AAQ28568

ID AAQ28568 standard; DNA; 3480 BP.

XX

AC AAQ28568;

XX

DT 25-MAR-2003 (updated)

DT 15-FEB-1993 (first entry)

XX Human thrombin receptor gene.

DE

XX Diagnosis; cardiovascular disease; wound healing; restenosis;

KW thrombosis; unstable angina treatment; myocardial infarction;

KW thrombotic; thromboembolytic stroke; ss.

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT CDS 225..1502

FT sig_peptide /*tag= a

FT mat_peptide 225..303

FT /*tag= b

FT 304..1502

FT /*tag= c

XX WO9214750-A1.

PN

XX 03-SEP-1992.

PD

Db 124 AAGACCGGCTCCCGACCCGAGAAAGTCAGGAGAGAGGGTGAAGCGGAGCAGCCCCGAGGC 183

QY 121 GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGGCCCGCGGGCTGC 180

Db 184 GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGGCCCGCGGGCTGC 243

QY 181 TGCTGGTGGCCGCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCGCCGACCCGGGCC 240

Db 244 TGCTGGTGGCCGCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCGCCGACCCGGGCC 303

QY 241 GCAGGCCAGAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCAATTCTTCAGGA 300

Db 304 GCAGGCCAGAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCAATTCTTCAGGA 363

QY 301 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA 360

Db 364 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA 423

QY 361 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCAT 420

Db 424 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCAT 483

QY 421 TCATCTCAGAAGATGCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 6

ACC50141

ID ACC50141 standard; cDNA; 3592 BP.

XX

AC ACC50141;

XX

DT 12-JUN-2003 (first entry)

XX

DE Breast cancer associated cDNA sequence SEQ ID NO:129.

XX

XX Human; breast cancer; cytostatic; gene therapy; gene; ss.

OS Homo sapiens.

XX WO2003004989-A2.

PN

XX PD 16-JAN-2003.

XX

PF 21-JUN-2002; 2002WO-US19669.

XX

PR 21-JUN-2001; 2001US-299887P.

PR 27-JUN-2001; 2001US-301572P.

PR 18-JUL-2001; 2001US-306501P.

PR 25-SEP-2001; 2001US-325002P.

PR 05-MAR-2002; 2002US-362585P.

PR 14-MAY-2002; 2002US-380391P.

XX

PA (MILL-) MILLENIUM PHARM INC.

XX

PI Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S, Mertens M;

PI Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;

PI Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX

DR WPI; 2003-210381/20.

DR P-PSDB; ABR47449.

XX

PT Breast cancer diagnosis or treatment by comparing the level of expression of a marker in a patient sample with that in the control

PT non-breast cancer sample -

XX

PS Claim 1; SEQ ID 129; 128pp; English.

XX

CC The present invention describes a method for assessing whether a patient is afflicted with breast cancer. The method comprises comparing the level of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and ABR47386 to ABR47632) in a patient sample and the normal level of expression of the marker in a control non-breast cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with breast cancer. The breast cancer associated sequences from the present invention have cytostatic activities and can be used in gene therapy. The method is useful for diagnosing and treating breast cancer.

CC N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 3592 BP; 950 A; 857 C; 837 G; 948 T; 0 other;

Query Match 100.0%; Score 548; DB 25; Length 3592;

Best Local Similarity 100.0%; Pred. No. 1.3e-138;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCTGCGCTCTGCTGCGCG 60

Db 184 CGCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCTGCGCTCTGCTGCGCG 243

QY 61 AAGACCGGCTCCCGACCCGACAGAGTCAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120

Db 244 AAGACCGGCTCCCGACCCGACAGAGTCAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 303

QY 121 GGGGAGCCTCCCGAGAGCGCGCGCAGAGCCCCGGGACAATGGGGCCCGCGGGTGC 180

Db 304 GGGGAGCCTCCCGAGAGCGCGCGCAGAGCCCCGGGACAATGGGGCCCGCGGGTGC 363

QY 181 TGCTGGTGGCCGCTGCTTCAGTCTGTGCGGCCCGCTGTGTCTGCCGACCCCGGCC 240

Db 364 TGCTGGTGGCCGCTGCTTCAGTCTGTGCGGCCCGCTGTGTCTGCCGACCCCGGCC 423

QY 241 GCAGGCCAGAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCAGGA 300

Db 424 GCAGGCCAGAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCAGGA 483

QY 301 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA 360

Db 484 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA 543

QY 361 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCAT 420

Db 544 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCAT 603

QY 421 TCATCTCAGAAGATGCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCAT 480

Db 604 TCATCTCAGAAGATGCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCAT 663

QY 481 CTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 664 CTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 723

QY 541 TCATCCTG 548

Db 724 TCATCCTG 731

RESULT 7

ABZ42767

ID ABZ42767 standard; DNA; 3592 BP.

XX

AC ABZ42767;

XX

DT 04-MAR-2003 (first entry)

XX DE Human thrombin receptor nucleotide SEQ ID NO:323.

XX KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;

KW G protein-coupled receptor modulator; antibody; immune-related disease;

KW growth-related disease; cell regenerative-related disease; AIDS; cancer;

KW immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;

KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;

KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;

KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;

KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

ulcer; gene; ds.

XX KW Homo sapiens.

OS WO200261087-A2.

XX PD 08-AUG-2002.

XX PF 19-DEC-2001; 2001WO-US50107.

XX PR 19-DEC-2000; 2000US-257144P.

XX PA (LIFE-) LIFESPAN BIOSCIENCES INC.

XX PI Burmer GC, Roush CL, Brown JP;

XX PI WPI; 2003-046718/04.

DR P-PSDB; ABP81919.

XX PT New isolated antigenic peptides e.g., for G protein-coupled receptors

PT (GPCR), useful for diagnosing and designing drugs for treating

PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,

PT cancer or autoimmune diseases -

XX PS Disclosure; Fig 1; 523pp; English.

XX CC The present invention describes antigenic peptides (I) comprising:

CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino

CC acids. Also described: (1) an assay for the detection of a particular

CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;

CC and (2) an isolated antibody having high specificity and high affinity

CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and

CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting

CC an antibody against a particular GPCR, and in the production of specific

CC antibodies. The peptides and antibodies are also useful for detecting the

CC presence or absence of corresponding GPCRs. The antigenic peptides for

CC GPCRs and antibodies are useful for diagnosing and designing drugs for

CC treating immune-related diseases, growth-related diseases, cell

CC regeneration-related disease, immunological-related cell proliferative

CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,

CC atherosclerosis, bacterial, fungal, protozoan or viral infections,

CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute

CC inflammation, allergies, Crohn's disease, diabetes, graft versus host

CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,

CC anxiety, depression, schizophrenia, dementia, mental retardation, memory

CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,

CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or

CC any other disorder in which GPCRs are involved. The antibodies may be

CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode

CC GPCR proteins given in ABP81675 to ABP82018, which are used in the

CC exemplification of the present invention.

XX SQ Sequence 3592 BP; 950 A; 857 C; 837 G; 948 T; 0 other;

Query Match 100.0%; Score 548; DB 25; Length 3592;

Best Local Similarity 100.0%; Pred. No. 1.3e-138;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX	DB	184	CGCCGAGGGTCGCTGGACCCCTGATCTTACCCGTGGGACCCCTGCGCTCTGCCTGCCGCG	243
DE	QY	61	AAGACCGGCTCCCGACCCCGCAGAGTTCAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC	120
XX	DB	244	AAGACCGGCTCCCGACCCCGCAGAGTTCAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC	303
KW	QY	121	GGGCGAGCCTCCCGGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGCGCGCGGGTGC	180
KW	DB	304	GGGCGAGCCTCCCGGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGCGCGCGGGTGC	363
KW	QY	181	TGCTGGTGGCGCGCTGCTTCACTCTGTGCGGCGCGCTGTTGTCTGCCCGCACCCGGGCC	240
KW	DB	364	TGCTGGTGGCGCGCTGCTTCACTCTGTGCGGCGCGCTGTTGTCTGCCCGCACCCGGGCC	423
KW	QY	241	GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCATTTTCTTCAGGA	300
OS	DB	424	GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCATTTTCTTCAGGA	483
XX	QY	301	ACCCCAATGATAAATATGAACCACTTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	360
PN	DB	484	ACCCCAATGATAAATATGAACCACTTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	543
XX	QY	361	CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT	420
XX	DB	544	CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT	603
PI	QY	421	TCATCTCAGAAGATGCCTCCGGATATTGACCAGCTCCTGGTGACACTCTTTGTGCCCAT	480
PI	DB	604	TCATCTCAGAAGATGCCTCCGGATATTGACCAGCTCCTGGTGACACTCTTTGTGCCCAT	663
DR	QY	481	CTGTGTACACCGGAGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	540
DR	DB	664	CTGTGTACACCGGAGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	723
XX	QY	541	TCATCCTG 548	
XX	DB	724	TCATCCTG 731	
CC	RESULT 8			
CC	ABK70889			
CC	ID	ABK70889 standard; cDNA; 1278 BP.		
CC	XX	ABK70889;		
CC	XX	15-JUL-2002 (first entry)		
CC	DE	Human cDNA encoding PAR1 type thrombin receptor.		
CC	XX	Human; ss; gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;		
CC	KW	inflammatory disease; cell proliferative disease.		
CC	OS	Homo sapiens.		
CC	PN	JP2002010784-A.		
CC	XX	15-JAN-2002.		
CC	XX	29-JUN-2000; 2000JP-0196514.		
CC	XX	29-JUN-2000; 2000JP-0196514.		
CC	XX	(TEIJ) TEIJIN LTD.		
CC	DR	WPI; 2002-321520/36.		
CC	DR	P-PSDB; ABG35300.		
CC	PT	An inhibitor of cell growth mediated by thrombin used to treat		
CC	PT	inflammatory and cell proliferative diseases -		
CC	XX	Claim 9; Page 26-27; 44pp; Japanese.		
CC	XX	The invention relates to a polypeptide or a compound which can inhibit		

CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
CC modified version).

XX Sequence 1278 BP; 276 A; 353 C; 284 G; 365 T; 0 other;

SQ Query Match 70.6%; Score 387; DB 24; Length 1278;
Best Local Similarity 100.0%; Pred. No. 4.8e-95;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ATGGGGCCGCGGCTGCTGGTGGCGCGCTGCTTCACTCTGTGGCGCCGCTGTTG 221
Db |||||
1 ATGGGGCCGCGGCTGCTGGTGGCGCGCTGCTTCACTCTGTGGCGCCGCTGTTG 60
QY 222 TCTGCCCGCACCCGCGCGCAGGCCAGAAATCAAAAAGCAACAATGCCACTTAGATCCC 281
Db |||||
61 TCTGCCCGCACCCGCGCGCAGGCCAGAAATCAAAAAGCAACAATGCCACTTAGATCCC 120
QY 282 CGGTCACTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAG 341
Db |||||
121 CGGTCACTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAG 180
QY 342 AAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTT 401
Db |||||
181 AAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTT 240
QY 402 CAAAAACAACCTTCTGCAATTCATCTCAGAAGATGCTCCGGATATTTGACACGCTCCTGG 461
Db |||||
241 CAAAAACAACCTTCTGCAATTCATCTCAGAAGATGCTCCGGATATTTGACACGCTCCTGG 300
QY 462 CTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAAC 521
Db |||||
301 CTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAAC 360
QY 522 ATCATGGCCATCGTTGTGTTTCATCCTG 548
Db |||||
361 ATCATGGCCATCGTTGTGTTTCATCCTG 387

RESULT 9

AAT62461

ID AAT62461 standard; DNA; 2910 BP.

XX AAT62461;

XX 25-MAR-2003 (updated)

DT 16-AUG-1997 (first entry)

XX ThR-GPA1 fusion gene.

DE ThR-GPA1 fusion gene.

XX G-protein coupled receptor; agonist; antagonist; assay;

KW G-alpha protein; Gpalp; GPA1 gene; thrombin receptor; ds.

XX Chimaeric Homo sapiens;

OS Chimaeric Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT CDS 1..2679

FT /tag= a

FT /product= ThR-GPA1 fusion

FT mRNA 1..39

FT /tag= b

FT /note= "STET nucleotides 535-573"

FT 40..1251

FT mRNA

FT /tag= c
FT /product= thrombin receptor
FT /note= "corresponds to nucleotides 288-1499 of
FT the thrombin receptor gene"
FT 1252..1260
FT /tag= d
FT /note= "STET nucleotides 1827..1835"
FT 1261..2910
FT /tag= e
FT /product= Gpalp
FT /note= "corresponds to nucleotides 202-1851 of
FT GPA1"
FT complement (42..62)
FT /tag= f
FT /note= "oligonucleotide g for thrombin receptor
FT gene amplification"
FT 1229..1251
FT /tag= g
FT /note= "oligonucleotide h for thrombin receptor
FT gene amplification"

WO9711159-A1.

27-MAR-1997.

20-SEP-1996; 96WO-US15203.

20-SEP-1995; 95US-0004023.

(HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.

Das OP, Mandell RB, Boulton TG, McMullen TW;

WPI; 1997-202868/18.

P-PSDB; AAW16314.

New yeast constructs for use in screening assays - which express a
fusion comprising a mammalian or fungal transmembrane receptor and
a yeast G-alpha protein

Example 2; Fig 12A-G; 109pp; English.

A gene fusion (AAT62461) codes for a polypeptide (AAW16314) comprising
the human thrombin receptor (ThR) covalently linked to yeast G-alpha
protein Gpalp (encoded by the GPA1 gene). It was obtd. by PCR
amplification (see also AAT62466-67) of the ThR gene and insertion
of the amplified sequence into pRMHT3, which contains the GPA1
gene. Thrombin-dependent activation of the yeast mating pathway
by the ThR-GPA1 fusion protein was observed in ste2 gpal- yeast
cells. The polypeptide is an example of novel fusion proteins
between eukaryotic heterotrimeric G-protein coupled receptors and
yeast G-alpha proteins that can be coupled to the pheromone-
induced signal transduction pathway of yeast and used in screening
assays to identify agonists or antagonists of the receptor.
(Updated on 25-MAR-2003 to correct PI field.)

Sequence 2910 BP; 868 A; 606 C; 628 G; 808 T; 0 other;

Query Match 59.9%; Score 328.4; DB 18; Length 2910;

Best Local Similarity 98.2%; Pred. No. 5.5e-79;

Matches 332; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 211 GCCCGCTGTGTCGCCCGCACCCCGCGCCGAGCCAGAAATCAAAAAGCAACAATGCCA 270

Db |||||

26 GCAATCTATTTATGCCCGCACCCCGCGCCGAGCCAGAAATCAAAAAGCAACAATGCCA 85

QY 271 CCTTAGATCCCGGTTCATTTCTCTCAGGAACCCCAATGATAAATATGAACCATTTGGG 330

Db |||||

86 CCTTAGATCCCGGTTCATTTCTCTCAGGAACCCCAATGATAAATATGAACCATTTGGG 145

QY 331 AGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAA 390

Db |||||

146 AGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAA 205

QY	391	GCAGTCCCTCTTTCAAAAACAACTTCCTGCAATTCATCTCAGAAGATGCCCTCCGGATATTGA	450
Dβ	206	GCAGTCCCTCTTTCAAAAACAACTTCCTGCAATTCATCTCAGAAGATGCCCTCCGGATATTGA	265
QY	451	CCAGCTCCTGGCTGACACTCTTTTGTCCTCATCTGTGTACACCGAGTGTTTGTAGTCAGCC	510
Dβ	266	CCAGCTCCTGGCTGACACTCTTTTGTCCTCATCTGTGTACACCGAGTGTTTGTAGTCAGCC	325
QY	511	TCCCACATAAACAATCATGGCCATCGTTGTGTTTCATCCTG	548
Dβ	326	TCCCACATAAACAATCATGGCCATCGTTGTGTTTCATCCTG	363

```

RESULT 10
AAF21433
ID   AAF21433 standard; DNA; 6546 BP.
XX
XX   AAF21433;
XX   AC
XX   DT      14-MAR-2001 (first entry)
XX
XX   DE      Human low adenosine antisense oligonucleotide related sequence #3000.

```

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

OS Homo sapiens.

PN WO200062736-A2.

PD 26-OCT-2000-

24-MAR-2000:

06-APR-1999: 99US-0127958.

XX PA (IIVC-) INTV EAST CAROLINA.

XX
XX
(NICH/ NICH C W):

FI NYCE OW;
XX

XX
XX
06/00000000-0007, TFM

PT trigger adenosine release
PT trigger adenosine release

FI cancers and respiratory obstructions -
yy

PS
XX
DISCLOSURE; Page 1418-1418; 1332pp; English

CC The present invention describes low ad-
CC clisynucleotides and compositions (T)

CC oligonucleotides the A is replaced by a
CC (T) can have respiratory bronchodilator an

The antitense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide

transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 6546 BP: 1603 A; 1691 C; 1500 G; 1751 T; 1 other;

```
Query Match          58.8%; Score 322.4; DB 21; Length 6546;
Best Local Similarity 99.7%; Pred. No. 3.2e-77;
Matches 323: Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 225 GCCCGCACCCGGCCGGCCAGGCAGAAATCAAAGCAACAAATGCCACCTTAGATCCCCGG 284

db 3362 GCACGCACCCGGCCGGCCAGGCAGAAATCAAAGCAACAAATGCCACCTTAGATCCCCGG 3421

QY 285 TCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATCAGGAGAAA 344

db 3422 TCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATCAGGAGAAA 3481

QY 345 AATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAA 404

D6 3482 AATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAA 3541

QY 405 AAACAACCTTCCTGCATTCACTCTCAGAAGATGCTCCGGATATTTGACCAAGCTCTCGGCTG 464

D β 3542 AAACAACCTTCCTGCATTCACTCTCAGAAGATGCTCCGGATATTTGACCAAGCTCTCGGCTG 3601

Qy 465 A C A C T C T T T G T C C C A T C T G T G T A C A C C G G A G T G T T G T A G T C A G C C T C C C A C T A A A C A T C 524

pB 3602 A C A C T C T T T G T C C C A T C T G T G T A C A C C G G A G T G T T G T A G T C A G C C T C C C A C T A A A C A T C 3661

525 ATGGCCATCGTTGTGTTTCATCCTG 548

3662 ATGGCCATCGTTGTGTTTCATCCTG 3685

RESULT 11

AAA35311

ID AAA35311 standard; DNA; 6546 BP.

AC AAA35311;

DT 28-JUL-2000 (first entry)

DE Human adenosine receptor x

KW Human; adenosine receptor; low adenosine antisense oligonucleotide

KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

PN W0200009525-A2

24-FEB-2000

03-AUG-1999:

W

PR 03-AUG-1998: 98US-0095212.

XX PA (IIV EC-) INIV EAST CAROLINA.

PI Nvce JW:

WPI: 2000-205971/18.

PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -

PS Disclosure: Page 1331-1333; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

sequence 6546 BP: 1603 A: 1692 C: 1500 G: 1750 T: 1 other; xx

```
Query Match      58.8%; Score 322.4; DB 21; Length 6546;
Best Local Similarity 99.7%; Pred. No. 3.2e-77;
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 225 GCGCGCACCGGGCCCGAGGCAGCAATCAAAAGCAACAAATGCCACCTTAGATCCCCG 284

Db 3362 GCACGCACCGGGCCCGAGGCAGCAATCAAAAGCAACAAATGCCACCTTAGATCCCCG 3421

3362 GCACGCACCCGGGCCCCGAGGCCAGAAATCAAAAGCAACAATGCCACCTTAGATCCCCGG 3421

Qy 285 TCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAGAA 344
|||
db 3422 TCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAGAA 3481

D_b 3422 TCATTCTCTCAGGAACCCCAATGATAAAATATGAACCAATTTTGGGAGGATGAGGAGAAA 3481

QY 345 AATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAA 404

D6 3482 AATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAA 3541

Db 3482 AATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCTCTTCAA 3541

Qy 405 AAACAACCTTCCTGCATTTCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTG 464
|||||
Db 3542 AAACAACCTTCCTGCATTTCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTG 3601
|||||

3543 AACACCTTCCTGCATTCACTCAGAAAGATGCCTCCGGATATTTGAACAGGTCCTGGCTG 3601

Qy 465 ACACCTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTCCCAATAACATC 524
|||
nb 3602 ACACCTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTCCCAATAACATC 3661

3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000

07 **E3E** **A T C G C C A A T C C T T G T T C A T C C M G** **548**

pk 3663 ATGCGCATCGTTCTCTTCAATCTG 3685

RESULT 12

AAAF21432

ID AAF21432 standard; DNA; 3182 BP.

AC AAF21432:

DT 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2999.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer: ss.

xx Homo sapiens.

AA WO200062736-A2

AA
PD
26-OCT-2000.

24-MAR-2000: 2000WO-US08020.

06-APR-1999: 99UIS-0127958.

XX
PA (IVEC-) INTV EAST CAROLINA

PA (UIEC-1) UNIV EAST
PA (NYCE/) NYCE J W

AA
PT
Nyce JW:

AX
DR WPI: 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not
 trigger adenosine receptors during metabolism, useful e.g. for treating
 cancers and respiratory obstructions -

PS Disclosure: Page 1415~1416: 1592pp: English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

XX

SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Query Match 58.6%; Score 321; DB 21; Length 3182;
Best Local Similarity 100.0%; Pred. No. 5.8e-77;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 CGCACCCGGCCGCGCAGGAGATCAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCA 287
Db 1 CGCACCCGGCCGCGCAGGAGATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCA 60

QY 288 TTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAAT 347
Db 61 TTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAAT 120

QY 348 GAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCTCTTCAAAAA 407
Db 121 GAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCTCTTCAAAAA 180

QY 408 CAACTTCCTGCATCTCAGAGATGCTCCGGATATTTTGACCAGTCCTCTGGCTGACA 467
Db 181 CAACTTCCTGCATCTCAGAGATGCTCCGGATATTTTGACCAGTCCTCTGGCTGACA 240

QY 468 CTCCTTGTCCCATCTGTGTACACCGGAGTCTTTGTAGTCAGCCCTCCCACTAAACATCATG 527
Db 241 CTCCTTGTCCCATCTGTGTACACCGGAGTCTTTGTAGTCAGCCCTCCCACTAAACATCATG 300

QY 528 GCCATCGTTGTGTTCATCCTG 548
Db 301 GCCATCGTTGTGTTCATCCTG 321

RESULT 13
AAA35310
ID AAA35310 standard; DNA; 3182 BP.
XX
AC AAA35310;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:184.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; cytotstatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 1330-1331; 1343pp; English.
XX

CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Query Match 58.6%; Score 321; DB 21; Length 3182;
Best Local Similarity 100.0%; Pred. No. 5.8e-77;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 CGCACCCGGCCGCGCAGGAGATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCA 287
Db 1 CGCACCCGGCCGCGCAGGAGATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCA 60

QY 288 TTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAAT 347
Db 61 TTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAAT 120

QY 348 GAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCTCTTCAAAAA 407
Db 121 GAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCTCTTCAAAAA 180

QY 408 CAACTTCCTGCATCTCAGAGATGCTCCGGATATTTTGACCAGTCCTCTGGCTGACA 467
Db 181 CAACTTCCTGCATCTCAGAGATGCTCCGGATATTTTGACCAGTCCTCTGGCTGACA 240

QY 468 CTCCTTGTCCCATCTGTGTACACCGGAGTCTTTGTAGTCAGCCCTCCCACTAAACATCATG 527
Db 241 CTCCTTGTCCCATCTGTGTACACCGGAGTCTTTGTAGTCAGCCCTCCCACTAAACATCATG 300

QY 528 GCCATCGTTGTGTTCATCCTG 548
Db 301 GCCATCGTTGTGTTCATCCTG 321

RESULT 14
AAQ73590
ID AAQ73590 standard; DNA; 1764 BP.
XX
AC AAQ73590;
XX
DT 25-MAR-2003 (updated)
DT 25-JUN-1995 (first entry)
XX
DE Fragment of the human thrombin receptor gene.
XX
KW TR; expression; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 20, 2003, 20:04:31 ; Search time 2381 Seconds
(without alignments)
9415.571 Million cell updates/sec

Title: US-09-744-679-7
Perfect score: 548
Sequence: 1 cgccgagggtcgcttgacc.....ccatcgttgtgttcacctg 548

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	548	100.0	3448	9	BC051909	BC051909 Homo sapi
2	548	100.0	3472	6	AR270765	AR270765 Sequence
3	548	100.0	3472	9	HUMTHRR	M62424 Human throm
4	548	100.0	3479	9	BC002464	BC002464 Homo sapi
5	548	100.0	3480	6	AR025392	AR025392 Sequence
6	548	100.0	3480	6	AR027730	AR027730 Sequence
7	548	100.0	3480	6	AR065265	AR065265 Sequence
8	548	100.0	3480	6	AR137610	AR137610 Sequence
9	548	100.0	3480	6	I74660	I74660 Sequence 21
10	548	100.0	3592	6	AX549038	AX549038 Sequence
11	441.6	80.6	3764	9	AF028727	AF028727 Papio ham
12	387	70.6	1278	6	BD107008	BD107008 Cell grow
13	387	70.6	1278	9	BT007279	BT007279 Homo sapi
14	306.2	55.9	70679	2	AC026151	AC026151 Homo sapi
15	302	55.1	24771	9	AF391809	AF391809 Homo sapi
16	302	55.1	110965	2	AC020927	AC020927 Homo sapi
17	302	55.1	143335	9	AC025188	AC025188 Homo sapi
18	282.4	51.5	1764	6	I49726	I49726 Sequence 12
19	249	45.4	88881	2	AC008933	AC008933 Homo sapi
20	246	44.9	2895	9	HSU63331	U63331 Human throm
21	240.8	43.9	1209	6	BD107006	BD107006 Cell grow
22	229	41.8	2878	9	AF054633	AF054633 Homo sapi
23	200.2	36.5	157274	2	AC143142	AC143142 Macaca mu
24	188.2	34.3	2855	10	CLTHRRCT	X61958 C.longicaud
25	184.2	33.6	1312	10	MUSTHRRCT	L03529 Mus musculu
26	183	33.4	3313	10	BC031516	BC031516 Mus muscu
27	171.2	31.2	3418	10	RATTRGPC	M81642 Rat G-prote
28	161	29.4	5457	9	HSU36755	U36755 Human throm
29	148.2	27.0	6436	6	AX345583	AX345583 Sequence
30	147	26.8	1116	6	BD107007	BD107007 Cell grow
31	146.6	26.8	6436	6	AX345582	AX345582 Sequence
32	137.6	25.1	182064	2	AC130634	AC130634 Rattus no
33	133	24.3	5620	10	MMTHREC02	U36757 Mus musculu
34	133	24.3	199227	2	AC110876	AC110876 Mus muscu
35	133	24.3	235097	2	AC110877	AC110877 Mus muscu
36	130.4	23.8	1164	10	CLU34047	U34047 Cricetulus
37	125.4	22.9	1300	6	AR212890	AR212890 Sequence
38	125.4	22.9	1300	6	AR212893	AR212893 Sequence
39	123.8	22.6	1300	6	AR212894	AR212894 Sequence
40	123.6	22.6	1298	6	AR212891	AR212891 Sequence
41	56	10.2	56	6	AR212873	AR212873 Sequence
42	55	10.0	55	6	AR212874	AR212874 Sequence
43	53	9.7	275	10	MMCF2R01	U55075 Mus musculu
44	53	9.7	4605	10	MMTHREC01	U36756 Mus musculu
45	52.6	9.6	1680	5	XLU09632	U09632 Xenopus lae

ALIGNMENTS

RESULT 1

BC051909

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC051909 3448 bp mRNA linear PRI 05-MAY-2003
Homo sapiens coagulation factor II (thrombin) receptor, mRNA (cDNA
clone MGC:60377 IMAGE:6172703), complete cds.

BC051909

BC051909.1 GI:30354671

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3448)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3448)
Strausberg,R.
Direct Submission
Submitted (01-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

FEATURES
source

gene

CDS

BASE COUNT 945 a 793 c 774 g 936 t
ORIGIN
Query Match 100.0%; Score 548; DB 9; Length 3448;
Best Local Similarity 100.0%; Pred. No. 1.1e-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGAGGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTCGCTCTGCTGCGCG 60
Db 26 CGCGAGGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTCGCTCTGCTGCGCG 85
QY 61 AAGACCGGCTCCCGACCCGACAGAAAGTCAGGAGAGAGGGTGAAGCGGAGAGAGAGGAGGC 120
Db 86 AAGACCGGCTCCCGACCCGACAGAAAGTCAGGAGAGAGGGTGAAGCGGAGAGAGGAGGC 145
QY 121 GGGGAGCCTCCCGGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGGCCGCGGGCTGC 180
Db 146 GGGGAGCCTCCCGGAGCAGCGCCGCGCAGAGCCCGGACAAATGGGGCCGCGGGCTGC 205
QY 181 TGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTGTCTGCCGACCCGGGCC 240
Db 206 TGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTGTCTGCCGACCCGGGCC 265
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGA 300
Db 266 GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGA 325
QY 301 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
Db 326 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 385
QY 361 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCTCTTCAAAAACAATCTTCTGCAT 420
Db 386 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCTCTTCAAAAACAATCTTCTGCAT 445
QY 421 TCATCTCAGAAGATGCCTCCGGATATTGACAGCTCCTGGCTGACACTCTTTGTCCCAT 480
Db 446 TCATCTCAGAAGATGCCTCCGGATATTGACAGCTCCTGGCTGACACTCTTTGTCCCAT 505
QY 481 CTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 506 CTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 565
QY 541 TCATCCTG 548
Db 566 TCATCCTG 573
RESULT 2
AR270765
LOCUS AR270765 3472 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1328 from patent US 6500938.
ACCESSION AR270765
VERSION AR270765.1 GI:29701999
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 3472)
AUTHORS Au-Young,J. and Seilhamer,J.J.
TITLE Composition for the detection of signaling pathway gene expression
JOURNAL Patent: US 6500938-A 1328 31-DEC-2002;
FEATURES Location/Qualifiers
source 1. .3472
/organism="unknown"
BASE COUNT 933 a 817 c 785 g 937 t
ORIGIN
Query Match 100.0%; Score 548; DB 6; Length 3472;
Best Local Similarity 100.0%; Pred. No. 1.1e-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 5 Row: p Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6031164.
FEATURES
source
1. .3479
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:2149"
/db_xref="taxon:9606"
/clone="MGC:1197 IMAGE:3343051"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC 16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
190. .1467
/codon_start=1
/product="coagulation factor II (thrombin) receptor"
/protein_id="AAH02464.1"
/db_xref="GI:12803297"
/translation="MGPRRLLLVAACFSLCGPLLSARTARRPESKATNATLDPRSFL
LRNPNDKYEPEFWEDEEKNEGLTEYRLVSINKSPLOKQLPAPFISEDASGYLTSSWLT
LFVPSVYTGTVFVSLPLNIMAJVFIKMKVKPAVVYMLHLATADVLFVSVLPFKIS
YFSGSDWQFGSELCRFVTAIFYCNMYASILLMTVISIDFLAVVYPMQSLSWRTLGR
ASFTCLAIWALAGVVPVLLKEQTIQVPLNITTTCHDVLNETLLEGYYAYYFSAFSA
VFFVPLIIISTVCYVSIIRCLSSSAVANRKSRLFLSAVFCIFICFGPTNVLLI
VHYSFLSHTSTTEAAYFAYLLCVSVSSISCCIDPLIYYASSECQRYVYSILCKESS
DPSSYNSGQLMASKMDTCSNLSNLSYKLLT"
BASE COUNT 969 a 792 c 778 g 940 t
ORIGIN
Query Match 100.0%; Score 548; DB 9; Length 3479;
Best Local Similarity 100.0%; Pred. No. 1.1e-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCCTGCGCTCTGCTGCCGCG 60
Db 29 CGCCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCCTGCGCTCTGCTGCCGCG 88
QY 61 AAGACCGGCTCCCCGAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 89 AAGACCGGCTCCCCGAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 148
QY 121 GGGGAGCCTCCCCGAGCAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 149 GGGGAGCCTCCCCGAGCAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 208
QY 181 TGCTGGTGGCGCGCTGCTTACGCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 240
Db 209 TGCTGGTGGCGCGCTGCTTACGCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 268
QY 241 GCAGGCGAGAATCAAAAGCAAAATGCCACCTTAGATCCCCCGGTCAATTTCTTCTCAGGA 300

Db 269 GCAGGCGAGAATCAAAAGCAAAATGCCACCTTAGATCCCCCGGTCAATTTCTTCTCAGGA 328
QY 301 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAGTGGTTAA 360
Db 329 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAGTGGTTAA 388
QY 361 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGCT 420
Db 389 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGCT 448
QY 421 TCATCTCAGAAGATGCTCCGGATATTGACCAAGTCTGGCTGACACTCTTTGTCCCAT 480
Db 449 TCATCTCAGAAGATGCTCCGGATATTGACCAAGTCTGGCTGACACTCTTTGTCCCAT 508
QY 481 CTGTGTACACCGGAGT 540
Db 509 CTGTGTACACCGGAGT 568
QY 541 TCATCCTG 548
Db 569 TCATCCTG 576
RESULT 5
AR025392
LOCUS AR025392 3480 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 219 from patent US 5798248
ACCESSION AR025392
VERSION AR025392.1 GI:3978020
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3480)
AUTHORS Coughlin, S.R. and Scarborough, R.M.
TITLE Recombinant mutant thrombin receptor and related pharmaceuticals
JOURNAL Patent: US 5798248-A 219 25-AUG-1998;
FEATURES
source 1. .3480
/organism="unknown"
BASE COUNT 941 a 817 c 785 g 937 t
ORIGIN
Query Match 100.0%; Score 548; DB 6; Length 3480;
Best Local Similarity 100.0%; Pred. No. 1.1e-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCCTGCGCTCTGCTGCCGCG 60
Db 64 CGCGAGGGTCGCTTGGACCTGATCTTACCCGTGGCACCCCTGCGCTCTGCTGCCGCG 123
QY 61 AAGACCGGCTCCCCGAGCAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 124 AAGACCGGCTCCCCGAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
QY 121 GGGGAGCCTCCCCGAGCAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 184 GGGGAGCCTCCCCGAGCAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
QY 181 TGCTGGTGGCGCGCTGCTTACGCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 240
Db 244 TGCTGGTGGCGCGCTGCTTACGCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 303
QY 241 GCAGGCGAGAATCAAAAGCAAAATGCCACCTTAGATCCCCCGGTCAATTTCTTCTCAGGA 300
Db 304 GCAGGCGAGAATCAAAAGCAAAATGCCACCTTAGATCCCCCGGTCAATTTCTTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAGTGGTTAA 360
Db 364 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGCT 420

Db 424 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 483

QY 421 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGGCTGACACTTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGGCTGACACTTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 6

AR027730

LOCUS AR027730 3480 bp DNA

DEFINITION Sequence 219 from patent US 5856448.

ACCESSION AR027730

VERSION AR027730.1 GI:5938550

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3480)

AUTHORS Coughlin,S.R.

TITLE Antibodies specifically reactive with thrombin receptor and its components

JOURNAL Patent: US 5856448-A 219 05-JAN-1999;

FEATURES Location/Qualifiers

source 1..3480

BASE COUNT 941 a 817 c 785 g 937 t

ORIGIN

Query Match 100.0%; Score 548; DB 6; Length 3480;

Best Local Similarity 100.0%; Pred. No. 1.1e-111;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCTGCCGCG 60

Db 64 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCTGCCGCG 123

QY 61 AAGACCGCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 120

Db 124 AAGACCGCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 183

QY 121 GGGGAGCCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 180

Db 184 GGGGAGCCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 243

QY 181 TGCTGGTGGCCGCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240

Db 244 TGCTGGTGGCCGCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 303

QY 241 GCAGGCCAGATCAAAAGCAACAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA 300

Db 304 GCAGGCCAGATCAAAAGCAACAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA 363

QY 301 ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAGAAATGAAGTGGGTAA 360

Db 364 ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAGAAATGAAGTGGGTAA 423

QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 420

Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 483

QY 421 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGCTGACACTTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGCTGACACTTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

Db 544 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 7

AR065265

LOCUS AR065265 3480 bp DNA

DEFINITION Sequence 219 from patent US 5849507.

ACCESSION AR065265

VERSION AR065265.1 GI:5995481

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3480)

AUTHORS Coughlin,S.R.

TITLE Methods to diagnose thrombosis by measuring activation peptide

JOURNAL Patent: US 5849507-A 219 15-DEC-1998;

FEATURES Location/Qualifiers

source 1..3480

BASE COUNT 941 a 817 c 785 g 937 t

ORIGIN

Query Match 100.0%; Score 548; DB 6; Length 3480;

Best Local Similarity 100.0%; Pred. No. 1.1e-111;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCTGCCGCG 60

Db 64 CGCCGAGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCTGCCGCG 123

QY 61 AAGACCGCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 120

Db 124 AAGACCGCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 183

QY 121 GGGGAGCCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 180

Db 184 GGGGAGCCTCCCGAGCAGCGCCGAGAGTCAAGAGAGAGGGTGAAGCGGAGAGCCCGAGGC 243

QY 181 TGCTGGTGGCCGCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240

Db 244 TGCTGGTGGCCGCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 303

QY 241 GCAGGCCAGATCAAAAGCAACAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA 300

Db 304 GCAGGCCAGATCAAAAGCAACAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA 363

QY 301 ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAGAAATGAAGTGGGTAA 360

Db 364 ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAGAAATGAAGTGGGTAA 423

QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 420

Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 483

QY 421 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGCTGACACTTTTGTCCCAT 480

Db 484 TCATCTCAGAAGATGCTCCGGATATTGACCAAGCTCCTGCTGACACTTTTGTCCCAT 543

QY 481 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540

Db 544 CTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603

QY 541 TCATCCTG 548

Db 604 TCATCCTG 611

RESULT 8
ARI137610
LOCUS ARI137610 3480 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 219 from patent US(6197541.)
ACCESSION ARI137610
VERSION ARI137610.1 GI:14479119
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3480)
AUTHORS Coughlin,S.R.
TITLE Recombinant thrombin receptors and assays using them
JOURNAL Patent: US 6197541-A 219 06-MAR-2001;
FEATURES Location/Qualifiers
source 1..3480
/organism="unknown"
BASE COUNT 941 a 817 c 785 g 937 t
ORIGIN
Query Match 100.0%; Score 548; DB 6; Length 3480;
Best Local Similarity 100.0%; Pred. No. 1.le-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGTCTGCTGCCGCG 60
Db 64 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGTCTGCTGCCGCG 123
QY 61 AAGACCCGCTCCCCGACCCGACGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 120
Db 124 AAGACCCGCTCCCCGACCCGACGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 183
QY 121 GGGGACGCTCCCCGAGCAGCGCGCGAGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 180
Db 184 GGGGACGCTCCCCGAGCAGCGCGCGAGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 243
QY 181 TGCTGGTGGCGCGCTGCTTCACTGTGTGCGGCCGCTGTGTCTGCGCCACCCGGGCC 240
Db 244 TGCTGGTGGCGCGCTGCTTCACTGTGTGCGGCCGCTGTGTCTGCGCCACCCGGGCC 303
QY 241 GCAGGCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCACTTCTCTCAGGA 300
Db 304 GCAGGCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCACTTCTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
Db 364 ACCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 483
QY 421 TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTGCTGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTGCTGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611
RESULT 9
I74660
LOCUS I74660 3480 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 219 from patent US 5688768.
ACCESSION I74660
VERSION I74660.1 GI:3010801
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3480)
AUTHORS Coughlin,S.R. and Scarborough,R.M.
TITLE Recombinant thrombin receptor and related pharmaceuticals
JOURNAL Patent: US 5688768-A 219 18-NOV-1997;
FEATURES Location/Qualifiers
source 1..3480
/organism="unknown"
BASE COUNT 941 a 817 c 785 g 937 t
ORIGIN
Query Match 100.0%; Score 548; DB 6; Length 3480;
Best Local Similarity 100.0%; Pred. No. 1.le-111;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGTCTGCTGCCGCG 60
Db 64 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGTCTGCTGCCGCG 123
QY 61 AAGACCCGCTCCCCGACCCGACGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 120
Db 124 AAGACCCGCTCCCCGACCCGACGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 183
QY 121 GGGGACGCTCCCCGAGCAGCGCGCGAGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 180
Db 184 GGGGACGCTCCCCGAGCAGCGCGCGAGAGTCAGGAGAGGGTGAAGCGGAGGAGCCGAGGC 243
QY 181 TGCTGGTGGCGCGCTGCTTCACTGTGTGCGGCCGCTGTGTCTGCGCCACCCGGGCC 240
Db 244 TGCTGGTGGCGCGCTGCTTCACTGTGTGCGGCCGCTGTGTCTGCGCCACCCGGGCC 303
QY 241 GCAGGCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCACTTCTCTCAGGA 300
Db 304 GCAGGCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCACTTCTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 360
Db 364 ACCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAGAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCTGTCAT 483
QY 421 TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTGCTGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTGCTGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGGAGTGTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611
RESULT 10
AX549038
LOCUS AX549038 3592 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 323 from Patent WO02061087.
ACCESSION AX549038
VERSION AX549038.1 GI:25813829
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Burner,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides

JOURNAL	Patent: WO 02061087-A 323 08-AUG-2002;									
FEATURES	Lifespan Biosciences, Inc. (US)									
source	Location/Qualifiers									
	1..3592									
	/organism="Homo sapiens"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:9606"									
BASE COUNT	950 a	857 c	837 g	948 t						
ORIGIN										
	Query Match	100.0%;	Score 548;	DB 6;	Length 3592;					
	Best Local Similarity	100.0%;	Pred. No. 1.1e-111;							
	Matches 548;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	CGCCGAGGGTCGCTTGGACCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG	60							
Db	184	CGCCGAGGGTCGCTTGGACCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG	243							
QY	61	AAGACCGGCTCCCCGACCCCGCAGAAATCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC	120							
Db	244	AAGACCGGCTCCCCGACCCCGCAGAAATCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC	303							
QY	121	GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCCGCGGCGGTGC	180							
Db	304	GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCCGCGGCGGTGC	363							
QY	181	TGCTGTTGGCCGCTGCTTCACTGTGTGCGGCCCGCTGTGTGTCTGCCCGCACCCGGGCC	240							
Db	364	TGCTGTTGGCCGCTGCTTCACTGTGTGCGGCCCGCTGTGTGTCTGCCCGCACCCGGGCC	423							
QY	241	GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA	300							
Db	424	GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCCGGTCAATTTCTTCTCAGGA	483							
QY	301	ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	360							
Db	484	ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	543							
QY	361	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAATTCCTGCAT	420							
Db	544	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAATTCCTGCAT	603							
QY	421	TCATCTCAGAAGATGCTCCGGATATTTGACCAGTCTCTGGCTGACACTCTTTGTCCCCT	480							
Db	604	TCATCTCAGAAGATGCTCCGGATATTTGACCAGTCTCTGGCTGACACTCTTTGTCCCCT	663							
QY	481	CTGTGTACACCGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	540							
Db	664	CTGTGTACACCGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	723							
QY	541	TCATCCTG	548							
Db	724	TCATCCTG	731							

RESULT 11	
AF028727	
LOCUS	AF028727 3764 bp mRNA linear PRI 20-NOV-2000
DEFINITION	Papio hamadryas thrombin receptor (bthr12) mRNA, complete cds.
ACCESSION	AF028727
VERSION	AF028727.1 GI:2605876
KEYWORDS	
SOURCE	Papio hamadryas (hamadryas baboon)
ORGANISM	Papio hamadryas
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
	Cercopithecinae; Papio.
	1 (bases 1 to 3764)
REFERENCE	Hayzer,D.J., Shoji,M. and Hanson,S.R.
AUTHORS	cdNAS encoding the baboon thrombin receptor indicate a primate
TITLE	transcription start site upstream of putative sites reported for
	the human gene
JOURNAL	Thromb. Res. 98 (2), 195-201 (2000)

```

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

20180190
10713321
2 (bases 1 to 3764)
Shoji,M., Hayzer,D.J. and Hanson,S.R.
Direct Submission
Submitted (07-OCT-1997) Hematologic Diseases Division, Building 1,
CDC/NCID/DHA, Rm. 1310, Mail Stop D-02, 1600 Clifton Road, Atlanta,
GA 30333, USA

FEATURES
source
1. .3764
/organism="Papio hamadryas"
/mol_type="mRNA"
/db_xref="taxon:9557"
1. .3764
/gene="bthrl2"
532. .1809
/gene="bthrl2"
/codon_start=1
/product="thrombin receptor"
/protein_id="AAB84191.1"
/db_xref="GI:2605877"
/translation="MGPRRLLVAACLCGPLLSARTRARRPASKATNATLDPRSFL
LRPNNDKYEPFEWDEEKNEESGLTEYRLVLSINKSPLQKPLPAFISEDAGYLTSSWLT
LFVPSVYTGFFVSLPVNIMAIVFYLKMKVKKPAVYMLHLATADVLFVSVLPFKIS
YILSGSDWQFGSELRCRFVTAAPFCNMYASILLMTVISIDRF LAVVYPMQSLSWRTLGR
ASFTCLAIWALAIAGVVPVLLKEQTIQVPGNLNITTCADVNETLLEGYYAYYFSAFSA
VFFVFLPIISTVCYVSIIRCLSSSTVANRSKSRALFSLAAVFCIFICFGPTNILLI
AHYFSLHTSTTEAAAYFAYLLCVCVSSISCCIDPLIYYVASSECQRVYVSYILCCRESS
DPSSSSGQLMAKMDTCSNLNNSIYKKLLT"

BASE COUNT      992 a   868 c   863 g   1041 t

```

Query Match	80.6%;	Score 441.6;	DB 9;	Length 3764;
Best Local Similarity	91.8%;	Pred. No. 5.6e-88;		
Matches 503;	Conservative 0;	Mismatches 34;	Indels 11;	Gaps 3;
QY	1	CGCCGAGGGTCGCTTGGACCCCTGATCTTACCCCGTGGGCACCCTGCGCTCTGCCTTGCCTGCGCG	60	
Db	382			
QY	61	AAGACCGGCTCCCCGACCCCGCAGAGTCAAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC	120	
Db	433	ACGACCGGCTGCCCGACCCAGCAGAAAGTTGGTAGAGAGGGTGA - CGGAGCAGCCCAAGGC	491	
QY	121	GGGGCAGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGACAATGGGSCCGCGCGGCTGC	180	
Db	492	GGGG - AACCTCTGGAGCAGCGCAGTCGGATCCCGGACAATGGGSCCGCGCGGCTGC	550	
QY	181	TGCTGTTGGCGCGCTGCTTCAGTCTGTGTGGCGCCCGCTGTTGTCTGCCCGCACCCCGGCGCC	240	
Db	551	TGCTGTTGGCGCGCTGCCCTCTGTCTGTGTGGCGCCCGCTGTTGTCCGACGACCCCGGCGCC	610	
QY	241	GCAGGCCAGAATCAAAAGCAACAATGCCACCTTAGATCCCCGGTCACTTCTTCTCAGGA	300	
Db	611	GCAGGCCAGCGTCAAAAGCAACAATGCCACCTTAGATCCCCGGTCACTTCTTCTCAGGA	670	
QY	301	ACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAGAGAAAATGAAAGTGGTTAA	360	
Db	671	ATCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAGAGAAAATGAAAGTGGTTAA	730	
QY	361	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCTAT	420	
Db	731	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCTAT	790	
QY	421	TCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCTGGGCTGACACTCTTTGTCCCAT	480	
Db	791	TCATCTCAGAAGATGCCTCCGGATATCTGACCAGCTCTGGGCTGACACTCTTTGTCCCAT	850	
QY	481	CTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	540	
Db	851	CTGTGTACACCGGGGTGTTGTGTAGTCAGCCTCCCGGTAAACATCATGGCCATCGTTGTGT	910	
QY	541	TCATCCTG	548	

Db 911 TCATCCTG 918

BD107008 1278 bp DNA linear PAT 18-SEP-2002

LOCUS Cell growth inhibitor.

DEFINITION BD107008

ACCESSION BD107008.1 GI:23201826

VERSION JP 2002010784-A/3.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1278)

Namiki,N., Suzuki,N., Tsunekawa,N., Kobayashi,S., Eguchi,H., Koike,Y. and Washimi,Y.

Cell growth inhibitor

Patent: JP 2002010784-A 3 15-JAN-2002;

TEIJIN LTD

OS Homo sapiens (human)

PN JP 2002010784-A/3

PD 15-JAN-2002

PF 29-JUN-2000 JP 2000196514

PI NAKO NAMIKI,NAOTAKA SUZUKI,NORIYUKI TSUNEKAWA,SHINJI PI KOBAYASHI.

PI HIROSHI EGUCHI,YUKINARI KOIKE,YOSHIHIKO WASHIMI PC C12N15/09,A61K38/00,A61P7/02,A61P9/10,A61P25/28,A61P29/00, PC A61P35/04,

PC C07K14/705,C07K16/28,C12P21/08,C12N15/00,A61K37/02 CC Cell growth inhibitor

FH Key

FT source

FT Location/Qualifiers

1..1278

/organism='Homo sapiens (human)'

Location/Qualifiers

1..1278

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 276 a 353 c 284 g 365 t

ORIGIN

Query Match 70.6%; Score 387; DB 6; Length 1278;

Best Local Similarity 100.0%; Pred. No. 7.9e-76;

Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ATGGGGCCGGCGGCTGCTGGTGGCGCGCTGCTTCAGTCTGTGGCGCCCGCTGTTG 221

Db 1 ATGGGGCCGGCGGCTGCTGGTGGCGCGCTGCTTCAGTCTGTGGCGCCCGCTGTTG 60

QY 222 TCTGCCCGCACCCGGCCCGCAGGCCAGAAATCAAAAAGCAACAAATGCCACCTTAGATCCC 281

Db 61 TCTGCCCGCACCCGGCCCGCAGGCCAGAAATCAAAAAGCAACAAATGCCACCTTAGATCCC 120

QY 282 CGGTCAATTTCTTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAG 341

Db 121 CGGTCAATTTCTTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAG 180

QY 342 AAAAAATGAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTT 401

Db 181 AAAAAATGAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTT 240

QY 402 CAAAAACAACCTTCCTGCATTCATCTCAGAGATGCCTCCGGATATTGACCGACTCCTGG 461

Db 241 CAAAAACAACCTTCCTGCATTCATCTCAGAGATGCCTCCGGATATTGACCGACTCCTGG 300

QY 462 CTGACACTCTTTGTCCTCATCTGTGTACACGGAGTGTTTGTAGTCAGCCTCCCACTAAAC 521

Db 301 CTGACACTCTTTGTCCTCATCTGTGTACACGGAGTGTTTGTAGTCAGCCTCCCACTAAAC 360

QY 522 ATCATGGCCATCGTTGTTCATCCTG 548

Db 361 ATCATGGCCATCGTTGTTCATCCTG 387

RESULT 13

BT007279 1278 bp mRNA linear PRI 13-MAY-2003

LOCUS Homo sapiens coagulation factor II (thrombin) receptor mRNA, complete cds.

DEFINITION BT007279

ACCESSION BT007279.1 GI:30583396

VERSION FLI_CDNA.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1278)

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.

Cloning of human full-length CDSs in BD Creator(TM) System Donor vector

Unpublished

2 (bases 1 to 1278)

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.

Direct Submission

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>.

Location/Qualifiers

1..1278

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="GH00863X1.0"

/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"

/lab_host="DH5alpha T1 resistant"

/note="Vector: pDNR-Dual"

1..1278

/codon_start=1

/product="coagulation factor II (thrombin) receptor"

/protein_id="AAP35943.1"

/db_xref="GI:30583397"

/translation="MGPRRLLVAAFCFLSGPLLSARTARRPESKATNATLDPKSFLLRNPNDRYEPFWEDEEKNEGLTEYRLVSINKSSPQQLPAFISEDASGYLTSSWLT LFPVSVYTGTVFVSLPLNMAIVFLLKMKVKPAVYVYMLHATADLVFVSVLPKIS YFSGSDWQFGSELRCFVTAAFYCNMYASILLMTVISIDRFVAVYPMQSLSWRTLGR ASFTCLAIWALAIAGVPLLLKEQTIQVPLNITTDHVLNETLLEGYAYYFSAFSA VFFVPLIISTVCYVSIIRCLSSAVANRKKSRALFLSAAVFCIFICFGPTNVLII VHYSFLSHTSTTEAAYFAYLLVCVSSISCCIDPLIYYVASSECQRVYVSILCKEBS DPSSYNSGQLMASKMDTCSSNLNNSIYKLLT"

BASE COUNT 276 a 352 c 284 g 366 t

ORIGIN

Query Match 70.6%; Score 387; DB 9; Length 1278;

Best Local Similarity 100.0%; Pred. No. 7.9e-76;

Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ATGGGGCCGGCGGCTGCTGTGGTGGCGCGCTGCTTCAGTCTGTGGCGCCCGCTGTTG 221

Db 1 ATGGGGCCGGCGGCTGCTGTGGTGGCGCGCTGCTTCAGTCTGTGGCGCCCGCTGTTG 60

* 26668 27483: contig of 816 bp in length
* 27484 27583: gap of 100 bp
* 27584 28412: contig of 829 bp in length
* 28413 28512: gap of 100 bp
* 28513 29344: contig of 832 bp in length
* 29345 29444: gap of 100 bp
* 29445 30271: contig of 827 bp in length
* 30272 30371: gap of 100 bp
* 30372 31143: contig of 772 bp in length
* 31144 31243: gap of 100 bp
* 31244 32053: contig of 810 bp in length
* 32054 32153: gap of 100 bp
* 32154 32962: contig of 809 bp in length
* 32963 33062: gap of 100 bp
* 33063 33883: contig of 821 bp in length
* 33884 34807: contig of 824 bp in length
* 34808 34907: gap of 100 bp
* 34908 35712: contig of 805 bp in length
* 35713 35812: gap of 100 bp
* 35813 36643: contig of 831 bp in length
* 36644 36743: gap of 100 bp
* 36744 37572: contig of 829 bp in length
* 37573 37672: gap of 100 bp
* 37674 38499: contig of 827 bp in length
* 38500 38599: gap of 100 bp
* 38600 39425: contig of 826 bp in length
* 39426 39525: gap of 100 bp
* 39526 40305: contig of 780 bp in length
* 40306 40405: gap of 100 bp
* 40406 41233: contig of 828 bp in length
* 41234 41333: gap of 100 bp
* 41334 42152: contig of 819 bp in length
* 42153 42252: gap of 100 bp
* 42253 43091: contig of 839 bp in length
* 43092 43191: gap of 100 bp
* 43192 44035: contig of 844 bp in length
* 44036 44135: gap of 100 bp
* 44136 44973: contig of 838 bp in length
* 44974 45073: gap of 100 bp
* 45074 45888: contig of 815 bp in length
* 45889 45988: gap of 100 bp
* 45989 46837: contig of 849 bp in length
* 46838 46937: gap of 100 bp
* 46938 47765: contig of 828 bp in length
* 47766 47865: gap of 100 bp
* 47866 48665: contig of 800 bp in length
* 48666 48765: gap of 100 bp
* 48766 49576: contig of 811 bp in length
* 49577 49676: gap of 100 bp
* 49677 50500: contig of 824 bp in length
* 50501 50600: gap of 100 bp
* 50601 51437: contig of 837 bp in length
* 51438 51537: gap of 100 bp
* 51538 52369: contig of 832 bp in length
* 52370 52469: gap of 100 bp
* 52470 53296: contig of 827 bp in length
* 53297 53396: gap of 100 bp
* 53397 54202: contig of 806 bp in length
* 54203 54302: gap of 100 bp
* 54303 55128: contig of 826 bp in length
* 55129 55228: gap of 100 bp
* 55229 56078: contig of 850 bp in length
* 56079 56178: gap of 100 bp
* 56179 57000: contig of 822 bp in length
* 57001 57100: gap of 100 bp
* 57101 57897: contig of 797 bp in length
* 57898 57997: gap of 100 bp
* 57998 58820: contig of 823 bp in length
* 58821 58920: gap of 100 bp
* 58921 59718: contig of 798 bp in length
* 59719 59818: gap of 100 bp
* 59819 60609: contig of 791 bp in length

* 60610 60709: gap of 100 bp
* 60710 61529: contig of 820 bp in length
* 61530 61629: gap of 100 bp
* 61630 62460: contig of 831 bp in length
* 62461 62560: gap of 100 bp
* 62561 63354: contig of 794 bp in length
* 63355 63454: gap of 100 bp
* 63455 64251: contig of 797 bp in length

Query Match 55.9%; Score 306.2; DB 2; Length 70679;
Best Local Similarity 69.6%; Pred. No. 9.1e-58;
Matches 332; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 72 CCCGAGCCCGCAGAGTTCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGCGGGGCGGCCTC 131
Db 51413 CCAGACAGATAAGAGTATGGGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 51472
QY 132 CCGGAGCAGCGCGCGCAGAGCCCGGAGCAATGGGGCGCGGGCTGCTGCTGTGGGCC 191
Db 51473 NNN 51532
QY 192 GCCTGCTTCAGTCTGTGCGGCGCGCTGTGTCTGCCGACCCCGCGGCGCAGGCAGAA 251
Db 51533 NNNNGCACACTTGTGCTGACCTGCAGGTTTCTCTAGNAGNATCCCCCTTCAGAA 51592
QY 252 TCAAAAGCAACAATGCCACCTTAGATCCCGGTCAATTTCTTCTCAGGAACCCCAATGAT 311
Db 51593 TCAAAAGCAACAATGCCACCTTAGATCCCGGTCAATTTCTTCTCAGGAACCCCAATGAT 51652
QY 312 AAATATGAACCAATTTGGGAGGATGAGGAGAGAAAATGAAAGTGGTTAACTGAATACAGA 371
Db 51653 AAATATGAACCAATTTGGGAGGATGAGGAGAGAAAATGAAAGTGGTTAACTGAATACAGA 51712
QY 372 TTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCTCATCTCAGAA 431
Db 51713 TTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCTCATCTCAGAA 51772
QY 432 GATGCTCCGATATTTGACCACTCTGCTGCTGACACTCTTGTCCCATCTGTGTACACC 491
Db 51773 GATGCTCCGATATTTGACCACTCTGCTGCTGACACTCTTGTCCCATCTGTGTACACC 51832
QY 492 GGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGCGCATCTGTGTTCATCCTG 548
Db 51833 GGAGTGTGTGTAGTCAGCCTCCCACTAAACATCATGCGCATCTGTGTTCATCCTG 51889

RESULT 15
AF391809 24771 bp DNA linear PRI 20-JUL-2001
LOCUS Homo sapiens coagulation factor II (thrombin) receptor (F2R) gene,
DEFINITION complete cds.
ACCESSION AF391809
VERSION AF391809.2 GI:14971463
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 24771)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REFERENCE 2 (bases 1 to 24771)
AUTHORS Rieder,M.J. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REMARK Sequence update by submitter
COMMENT On Jul 20, 2001 this sequence version replaced gi:14583134.
To cite this work please use: SeattleSNPs. NHLBI Program for
Genomic Applications, UW-FHCRC, Seattle, WA (URL:

variation /rpt_family="AluSx"
/rpt_type=dispersed
5159
/gene="F2R"
/frequency="0.02"
/replace="T"
5210
variation /gene="F2R"
/frequency="0.01"
/replace="C"
5276
variation /gene="F2R"
/frequency="0.03"
/replace="A"
5428
variation /gene="F2R"
/frequency="0.01"
/replace="A"
5472
variation /gene="F2R"
/frequency="0.40"
/replace="A"
5480
variation /gene="F2R"
/frequency="0.01"
/replace="A"
5515
variation /gene="F2R"
/frequency="0.01"

Query Match	55.1%;	Score 302;	DB 9;	Length 24771;
Best Local Similarity	100.0%;	Pred. No. 7.6e-57;		
Matches 302;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	247	CAGAAATCAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCA	306	
Db	18153	CAGAAATCAAAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCA	18212	
QY	307	ATGATAAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAAT	366	
Db	18213	ATGATAAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAAT	18272	
QY	367	ACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAACAACCTTCCTGCATTCACTCT	426	
Db	18273	ACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAACAACCTTCCTGCATTCACTCT	18332	
QY	427	CAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGT	486	
Db	18333	CAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGT	18392	
QY	487	ACACCGGAGTGTTTGTAGTCAGCCTCCCACATAACATCATGGCCATCGTTGTGTTTCATCC	546	
Db	18393	ACACCGGAGTGTTTGTAGTCAGCCTCCCACATAACATCATGGCCATCGTTGTGTTTCATCC	18452	
QY	547	TG 548		
Db	18453	TG 18454		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2003, 21:53:50 ; Search time 1751 Seconds
(without alignments)
7606.427 Million cell updates/sec

Title: US-09-744-679-7

Perfect score: 548

Sequence: 1 cgccgagggtcgcttgacc.....ccatgtgtgttcacctcg 548

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	548	100.0	888	13	BQ440574
2	536	97.8	699	10	BE258615
3	505.2	92.2	739	10	BG473471
4	251.2	45.8	1072	13	BU173240

5	197	35.9	496	11	BC016059	BC016059 Homo sapi
6	184	33.6	457	14	CA892151	CA892151 B0170E01-
7	184	33.6	575	14	CA892959	CA892959 B0175G06-
8	184	33.6	997	13	BU839951	BU839951 AGENCOURT
9	183	33.4	519	4	BX512139	BX512139 RZPD Mus
10	183	33.4	621	10	BB662299	BB662299 BB662299
11	183	33.4	642	14	BY754069	BY754069 BY754069
12	183	33.4	658	14	BY753574	BY753574 BY753574
13	183	33.4	674	12	BI685426	BI685426 603308867
14	183	33.4	676	14	BY754521	BY754521 BY754521
15	183	33.4	942	13	BQ896389	BQ896389 AGENCOURT
16	183	33.4	3338	11	AK085990	AK085990 Mus muscu
17	182.4	33.3	461	14	CA566568	CA566568 K0402H07-
18	182.4	33.3	492	14	CD554710	CD554710 B0381D04-
19	182.4	33.3	949	13	BQ923464	BQ923464 AGENCOURT
20	182	33.2	705	14	BY737790	BY737790 BY737790
21	181.4	33.1	513	14	CD554324	CD554324 B0374E01-
22	180.8	33.0	584	12	BI738768	BI738768 603362614
23	175.8	32.1	688	9	AA882290	AA882290 vx43b01.r
24	169.2	30.9	1031	13	BU516402	BU516402 AGENCOURT
25	168.4	30.7	456	14	CA877850	CA877850 K0960H08-
26	166	30.3	445	14	CA880312	CA880312 K0981H11-
27	165.6	30.2	672	10	BE911999	BE911999 601662765
28	164.2	30.0	709	13	BU611590	BU611590 UI-M-FIO-
29	162.2	29.6	763	12	BI666195	BI666195 603287191
30	159.2	29.1	467	10	BB862683	BB862683 BB862683
31	157.6	28.8	466	10	BB852109	BB852109 BB852109
32	152.2	27.8	410	13	BY228821	BY228821 BY228821
33	152	27.7	1429	13	BQ926183	BQ926183 AGENCOURT
34	145.8	26.6	405	10	BF116952	BF116952 uy92h02.y
35	136.2	24.9	403	13	BY301816	BY301816 BY301816
36	135.8	24.8	374	14	CA876551	CA876551 K0949C07-
37	135.8	24.8	461	14	CA878584	CA878584 K0968H09-
38	135.2	24.7	411	13	BY109295	BY109295 BY109295
39	133.6	24.4	413	13	BY228431	BY228431 BY228431
40	133	24.3	557	10	BG067569	BG067569 H3055G12-
41	132.6	24.2	412	13	BY228469	BY228469 BY228469
42	132	24.1	554	14	C87335	C87335 C87335 Mous
43	132	24.1	588	14	C86764	C86764 C86764 Mous
44	130.8	23.9	399	13	BY228641	BY228641 BY228641
45	128	23.4	504	9	AA690768	AA690768 vt25e01.r

ALIGNMENTS

RESULT 1
BQ440574
LOCUS
DEFINITION AGENCOURT 7907353 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154861
888 bp mRNA linear EST 24-MAY-2002
5', mRNA Sequence.
ACCESSION BQ440574
VERSION BQ440574.1 GI:21179650
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 888)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13496 row: j column: 14
High quality sequence stop: 654.

FEATURES		Location/Qualifiers			
source		1. .888			
		/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:6154861"			
		/tissue_type="retinoblastoma"			
		/lab_host="DH10B (phage-resistant)"			
		/clone_lib="NIH_MGC_67"			
		/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."			
BASE COUNT		181 a	263 c	237 g	207 t
ORIGIN					
Query Match		100.0%; Score 548; DB 13; Length 888;			
Best Local Similarity		100.0%; Pred. No. 3.1e-131;			
Matches 548; Conservative		0; Mismatches	0; Indels	0; Gaps	0;
QY	1	CGCCGAGGGTCGCTTGACCTGATCTTACCCGTGGGCACCTCGCTCTGCTGCGCG	60		
Db	25	CGCCGAGGGTCGCTTGACCTGATCTTACCCGTGGGCACCTCGCTCTGCTGCGCG	84		
QY	61	AAGACCGGCTCCCGACCCGAGAGTCAAGTCAGGAGAGGGTGAAGCGGAGCAGCCGAGGC	120		
Db	85	AAGACCGGCTCCCGACCCGAGAGTCAAGTCAGGAGAGGGTGAAGCGGAGCAGCCGAGGC	144		
QY	121	GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCGCGCGCTGC	180		
Db	145	GGGGCAGCCTCCCGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCGCGCGCTGC	204		
QY	181	TGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCGCGCTGTTGTCTGCCCGCACCCGGGCC	240		
Db	205	TGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCGCGCTGTTGTCTGCCCGCACCCGGGCC	264		
QY	241	GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCGGTCAATTTCTTCAGGA	300		
Db	265	GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCGGTCAATTTCTTCAGGA	324		
QY	301	ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	360		
Db	325	ACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAA	384		
QY	361	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT	420		
Db	385	CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT	444		
QY	421	TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTCTGGCTGACACTCTTTGTCCCAT	480		
Db	445	TCATCTCAGAAGATGCTCCGGATATTTGACCAAGTCTCTGGCTGACACTCTTTGTCCCAT	504		
QY	481	CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	540		
Db	505	CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT	564		
QY	541	TCATCCTG	548		
Db	565	TCATCCTG	572		
RESULT 2					
BE258615					
LOCUS					
DEFINITION BE258615 699 bp mRNA linear EST 13-JUL-2000					
601106874F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3343051 5',					
mRNA sequence.					
ACCESSION BE258615					
VERSION BE258615.1 GI:9129107					
KEYWORDS EST.					
SOURCE Homo sapiens (human)					
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					

REFERENCE	1 (bases 1 to 699)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCM127 row: c column: 20.				
FEATURES	Location/Qualifiers				
source	1..699				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:3343051"				
	/tissue_type="retinoblastoma"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_lib="NIH_MGC_16"				
	/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally-cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
BASE COUNT	153 a	202 c	194 g	149 t	1 others
ORIGIN					
Query Match	97.8%; Score 536; DB 10; Length 699;				
Best Local Similarity	99.6%; Pred. No. 3.7e-128;				
Matches 547; Conservative	0;	Mismatches	1;	Indels	1; Gaps 1;
QY	1	CGCCGAGGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCTCGCTCGCTGCGCGG	60		
Db	51	CGCCGAGGGTCGCTTGACCCCTGATCTTACCCGTGGGCACCTCGCTCGCTGCGCGG	110		
QY	61	AAGACCGGCTCCCGACCCGACAGTCAAGTCAGGAGAGGGTGAACGGAGAGCCCGAGGC	119		
Db	111	AAGACCGGCTCCCGACCCGACAGTCAAGTCAGGAGAGGGTGAACGGAGAGCCCGAGGC	170		
QY	120	CGGGGAGCCTCCCGAGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCCGCGCGGCTG	179		
Db	171	CGGGGAGCCTCCCGAGAGCAGCGCCGCGCAGAGCCCGGACAATGGGGCCGCGCGGCTG	230		
QY	180	CTGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCGCGCTGTTGTCTGCCGACCCGGGCC	239		
Db	231	CTGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCGCGCTGTTGTCTGCCGACCCGGGCC	290		
QY	240	CGCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGG	299		
Db	291	CGCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCAATTTCTTCAGG	350		
QY	300	AACCCCAATGATAAATATGAACCAATTTTGGAGGATGAGGAGAAAAATGAAAGTGGTTA	359		
Db	351	AACCCCAATGATAAATATGAACCAATTTTGGAGGATGAGGAGAAAAATGAAAGTGGTTA	410		
QY	360	ACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCCTGCA	419		
Db	411	ACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAAACAACCTTCCTGCA	470		
QY	420	TTCATCTCAGAAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCA	479		
Db	471	TTCATCTCAGAAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCA	530		
QY	480	TCTGTGTACACCGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTG	539		
Db	531	TCTGTGTACACCGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTG	590		

QY 540 TTCAATCCTG 548
|||||
Db 591 TTCAATCCTG 599

RESULT 3
BG473471
LOCUS
DEFINITION BG473471 739 bp mRNA linear EST 21-MAR-2001
602516274F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4647883 5',
mRNA sequence.

ACCESSION BG473471
VERSION BG473471.1 GI:13405746
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 739)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1425 row: C column: 20
High quality sequence stop: 718.
Location/Qualifiers
1. .739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4647883"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_16"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
162 a 207 c 200 g 170 t

BASE COUNT
ORIGIN
Query Match 92.2%; Score 505.2; DB 10; Length 739;
Best Local Similarity 99.1%; Pred. No. 3.7e-120;
Matches 529; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 16 GGACCTGATCTTACCCGTGGGCACCCCTGCGTCTGCCTGCCGCGAAGACCCGGCTCCCCG 75
|||||
Db 2 GGACCTGATCTTACCCGTGGGCACCCCTGCGTCTGCCTGCCGCGAAGACCCGGCTCCCCG 61
|||||

QY 76 ACCCGCAGAACTCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGGGGGCGAGCCTCCCCG 135
|||||
Db 62 ACCCGCAGAACTCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGGGGGCGAGCCTCCCCG 121
|||||

QY 136 AGCAGCGCCGCGCAGACCCGCGACAATGGGCGCGCGCGGTGCTGCTGTGGCGCGCT 195
|||||
Db 122 AGCAGCGCCGCGCAGACCCGCGACAATGGGCGCGCGCGGTGCTGCTGTGGCGCGCT 181
|||||

QY 196 GCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCACCCGGCCCGCAGGCCAGAAATCAA 255
|||||
Db 182 GCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCACCCGGCCCGCAGGCCAGAAATCAA 241
|||||

QY 256 AAGCAACAAATGCCACCTTAGATCCCCGGTCAATTTCTTCAGGAACCCCAATGATAAT 315
|||||

Db 242 AAGCAACAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAAT 301
|||||
QY 316 ATGAACCAATTTGGGAGGATGAGGAGAGAAAAATGAAAGTGGTTAACTGAATACAGATTAG 375
|||||
Db 302 ATGAACCAATTTGGGAGGATGAGGAGAGAAAAATGAAAGTGGGTAAAGTGAATACAGATTAG 361
|||||

QY 376 TCTCATCAATAAAAGCAGTCCTCTTCAAAAAACAATTCC-TGCATTCAATCTCAGAAGAT 434
|||||
Db 362 TCTCATCAGTAAAGCAGTCCTCTTCAAAAAACAATTCC-TGCATTCAATCTCAGAAGAT 421
|||||

QY 435 GCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGA 494
|||||
Db 422 GCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTGGTCCCCTGTGTGTACACCGGA 481
|||||

QY 495 GTGTTGTAGTCAGCCTCCCCTCACTAAACATCATGGCCATCGTTGTGTTTCATCCTG 548
|||||
Db 482 GTG-TTGTAGTCAGCCTCCCCTCACTAAACATCATGGCCATCGTTGTGTTTCATCCTG 534
|||||

RESULT 4
BUI73240
LOCUS
DEFINITION BUI73240 1072 bp mRNA linear EST 04-SEP-2002
AGENCOURT 7939641 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6172703
5', mRNA sequence.

ACCESSION BUI73240
VERSION BUI73240.1 GI:22687224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1072)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13543 row: a column: 24
High quality sequence stop: 149.
Location/Qualifiers
1. .1072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6172703"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 247 a 320 c 289 g 211 t 5 others
ORIGIN
Query Match 45.8%; Score 251.2; DB 13; Length 1072;
Best Local Similarity 92.2%; Pred. No. 3e-54;
Matches 273; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 60
|||||
Db 24 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 83
|||||

QY 61 AAGACCGGCTCCCCGACCCCGCAGAAAGTCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
|||||
Db 84 AAGACCGGCTCCCCGACCCCGCAGAAAGTCAGGAGAGAGGGTGAAGCGGAGCAGCCCGAGGC 143
|||||

QY 121 GGGCAGCCTCCCGAGCAGCGCCGCGAGAGCCCGGGACAATGGGGCCGCGCGGCTGC 180
|||||
Db 144 GGGCAGCCTCCCGAGCAGCGCCGCGAGAGCCCGGGACNATGGNGCGCGGCTGC 203
|||||
QY 181 TGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCA-CCCGGGCC 239
|||||
Db 204 TGCTGGNTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCACCCCGGGCC 263
|||||
QY 240 CGCAGGCCAGAAATCAAAAGCAACAATGCCACCTTAGATCCCGGTCAATTCTTCT 295
|||||
Db 264 GGCAAGCCANAATCAAAAGCAACAATGCCACCCCTAAATCCCGGGTCATTCT 319
|||||

RESULT 5
BC016059
LOCUS BC016059 496 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, Similar to coagulation factor II (thrombin) receptor,
clone IMAGE:4849569, mRNA.

ACCESSION BC016059
VERSION BC016059.1 GI:16359189
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 34 Row: a Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6031164
This clone has the following problem: retained intron.

FEATURES
Location/Qualifiers
1..496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4849569"
/tissue_type="Blood, natural killer cell"
/clone_lib="NIH MGC_106"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT 130 a 144 c 146 g 76 t
ORIGIN

Query Match 35.9%; Score 197; DB 11; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.6e-40;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CTGCGCGAAGACCGGCTCCCGACCCGAGAGTCAAGAGAGAGGGTGAAAGCGAGCAG 112
|||||
Db 1 CTGCGCGGAAGACCGGCTCCCGACCCGAGAGTCAAGAGAGAGGGTGAAAGCGAGCAG 60
|||||
QY 113 CCGAGCGGGGCGAGCCTCCCGGAGCAGCGCGGAGACCCGGGACAAATGGGCGCGCG 172
|||||
Db 61 CCGAGCGGGGCGAGCCTCCCGGAGCAGCGCGCGGAGACCCGGGACAAATGGGCGCGCG 120
|||||
QY 173 GCGGCTGCTGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCAC 232
|||||
Db 121 GCGGCTGCTGCTGGTGGCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCTGCCCGCAC 180
|||||
QY 233 CCGGGCCCCGAGGCCAG 249
|||||
Db 181 CCGGGCCCCGAGGCCAG 197
|||||

RESULT 6
CA892151
LOCUS CA892151 457 bp mRNA linear EST 20-DEC-2002
DEFINITION B0170E01-5N NIA Mouse Neural Stem Cell (Differentiated) cDNA
Library (Long) Mus musculus cDNA clone NIA:B0170E01 IMAGE:30100560
5', mRNA sequence.

ACCESSION CA892151
VERSION CA892151.1 GI:27343700
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Aiba,K., Vescovi,A.L.
and Ko,M.S.H.
TITLE Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated)
cDNA Library (Long)
JOURNAL Unpublished
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0170 row: E column: 01
Seq primer: -21M13 Reverse
High quality sequence stop: 457
POLYA=No.

FEATURES
Location/Qualifiers
1..457
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="niaEST:B0170E01-5N"
/db_xref="taxon:10090"
/clone="NIA:B0170E01 IMAGE:30100560"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="NIA Mouse Neural Stem Cell (Differentiated)
cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTTCTAGATCGGAGCGCGCCCTTTTTTT-3'] from
2.0 Microgram of total RNA, treated with T4 DNA polymerase
, and purified by ethanol-precipitation. The cDNAs were
ligated to Lone-linker LL-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by


```

DEFINITION  AGENCOURT_8947737 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6329922
5', mRNA sequence.
ACCESSION  BU839951
VERSION    BU839951.1 GI:24024346
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 997)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13783 row: h column: 19
            High quality sequence stop: 640.
            Location/Qualifiers
                1..997
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:6329922"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_130"
                /note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccd;
                Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
                Primer: Oligo dT. Average insert size 1.95 kb.
                Constructed by ResGen, Invitrogen Corp. Note: this is a
                NIH_MGC Library."
BASE COUNT  178 a 318 c 261 g 238 t      2 others
ORIGIN
Query Match      33.6%; Score 184; DB 13; Length 997;
Best Local Similarity 67.0%; Pred. No. 7.9e-37;
Matches 284; Conservative 0; Mismatches 125; Indels 15; Gaps 1;

QY  140 GCGCGCGCAGAGCCCGGACAAATGGGGCGCGCGCTGCTGCTGGTGGCGCGCTT 199
Db  140 GCGCGCGCGCAGCGCTTGGGAAATGGGGCGCGCGCTTGTGCTGCTGCGCGCTT 63

QY  200 CAGTCTGTGCGCGCGCGCTGTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 259
Db  64 CAGCTGTGCGGTCCCTTGTGCTGCTTCCCGCGCTTATGAGCCAGGAGATCAGAGAG 123

QY  260 AACAAATGCCACCTTAGATCCCGCGTCAATTTCTTCTCAGGAACCCCAATGATAATGA 319
Db  124 GACAGATGCTACGTGTAACCCCGCTCATTTCTTAAGGAATCCCAATGATAATGATT 183

QY  320 ACCATT-----TTGGGAGGATGAGGAGAAATGAAGTGGTTAACTGA 364
Db  184 TGAACCTGTCCTCCCTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 243

QY  365 ATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACTTCCTGCTTCAT 424
Db  244 GGGTAGGCGAGTCTACTTAATAATAGAGCTCCCTTCTCACACGCGCGCTCCTTCCCTTCAT 303

QY  425 CTCAGAGATGCTCCCGGATTAATGACCGAGTCTCTGGCTGACACTCTTGTCCCATCTGT 484
Db  304 CTCCGAGGACGCTCCCGGATATCTGACCGAGCCCTGCTGACGCTCTTCTGATGCGCTCCGT 363

QY  485 GTACACCGAGTGTGTTAGTCAGCGCTCCCAATAACATCATGGCCATGCTGTGTTTCAT 544
Db  364 GTACACGATTGTGTTTCAATGTCAGCGCTTCTCTGTAACGCTCCTGGCCATGCGAGTGTTCGT 423

```

```

QY  545 CCTG 548
    |||
Db  424 CTTG 427

RESULT 9
BX512139
ID  BX512139 standard; RNA; EST; 519 BP.
XX
AC  BX512139;
XX
SV  BX512139.1
XX
DT  27-MAY-2003 (Rel. 75, Created)
DT  27-MAY-2003 (Rel. 75, Last updated, Version 1)
XX
DE  RZPD Mus musculus cDNA clone IMAGp952C114 = IMAGE:337258 5' EST.
XX
KW  EST; expressed sequence tag.
XX
OS  Mus musculus (house mouse)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX
RN  [1]
RP  1-519
RA  Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
RA  Korn B.;
RT  ;
RL  Submitted (28-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
RL  Feld 580, D-69120 Heidelberg, Germany
XX
RZPD; IMAGp952C114.
CC  RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
CC  Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
CC  http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981
CC  Contact: Ina Rolfs
CC  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
CC  Heubnerweg 6, D-14059 Berlin, Germany
CC  Tel: +49 30 32639 101
CC  Fax: +49 30 32639 111
CC  www.rzpd.de
CC  This clone is available royalty-free from RZPD;
CC  contact RZPD (clone@rzpd.de) for further information.
CC  Seq primer: T7, Primer sequence: TAATACGACTCACTATAGGG
XX
Key  Location/Qualifiers
FH  source
FT  1..519
FT  /db_xref="taxon:10090"
FT  /note="1st strand cDNA was primed with a Not I - oligo(dT)
FT  primer [5'
FT  TGTACCAATCTGAAGTGGAGCGCGCGCATGTTTTTTTTTTTTTTT 3'], on
FT  total mouse RNA [provided by Minoru Ko, Wayne State Univ.];
FT  double-stranded cDNA was ligated to EcoRI adaptors
FT  5'-AATTCGGCAGCGG-3' and 5'-CCTCGTGGCG-3' (Pharmacia),
FT  digested with NotI and cloned into the NotI and EcoRI sites
FT  of the pT7T3D-PacI vector. Library went through one round
FT  of normalization, and was constructed by Bento Soares and
FT  M.Fatima Bonaudo."
FT  /organism="Mus musculus"
FT  /clone="IMAGp952C114"
FT  /clone_lib="Soares 19.5dpc p3NMF19.5"
FT  /dev_stage="19.5dpc total fetus"
FT  /lab_host="DH10B"
XX
SQ  Sequence 519 BP; 101 A; 160 C; 149 G; 109 T; 0 other;

Query Match      33.4%; Score 183; DB 4; Length 519;
Best Local Similarity 65.3%; Pred. No. 1.1e-36;
Matches 292; Conservative 0; Mismatches 140; Indels 15; Gaps 1;

```


cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 cDNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11876 row: e column: 17
 High quality sequence stop: 672.
 Location/Qualifiers

```

1. 674
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5344888"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
COUNT 126 a 212 c 188 g 148 t
ry Match 33.4%; Score 183; DB 12; Length 674;
t Local Similarity 65.3%; Pred. No. 1.2e-36;
ches 292; Conservative 0; Mismatches 140; Indels 15; Gaps 1;

117 AGGCGGGGAGGCTCCCGGAGCAGCGCGCGGACAAATGGGGCCGCGCGG 176
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
37 AGTAGCGGACGGCGGGCGCGCGCGGACAAATGGGGCCGCGCGG 96
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

177 CTGTGCTGTGGCGCGCTGCTTCAGTCTGTGCGCGCGCGCTGTGCTGCTGCCCGCACCGG 236
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 TTGCTGATGCTGCGCGCTCGCGCTCAGCTGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 156
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

237 GCGCGGAGCGGAGGATCAAAAGCAACAAATGCCAGTCCCGGTCATTTCTCTC 296
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 ATGAGCCAGCCAGAAATCAGAGAGGACAGATGCTACGGTGAACCCCGCTCATTTCTTA 216
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

297 AGGAACCCCAATGATAAATATGAACCAATTTGAACTGGTCCCGCTGGGATGAGGAGGAG 341
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 AGGAATCCCAATGATAAATATGAACCAATTTGAACTGGTCCCGCTGGGATGAGGAGGAG 276
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

342 AATAATGAAGTGGGTTAACTGAATCAGATTAAGTCCATCAATAAAGAGTCTCTT 401
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
277 AATAATGAAGGCTCCTGCTGGAGGGTAGGGCAGTCTACTTAATATAGCTCCCTCT 336
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

402 CAATAACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 CACAGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 396
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

462 CTGACACTCTTTGCTCCCATCTGTGTACACCGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTT 521
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
397 CTGACGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

522 ATCATGCGCCATCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 548
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
457 GTCTGCGCCATCGCAGTGTTCGTCCTG 483
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
 BY754521
 LOCUS
 DEFINITION
 musculus cDNA clone 142004K05 5', mRNA sequence.
 BY754521

musculus cDNA clone 142004K05 5', mRNA sequence.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 676)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
 Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chothia, C., Corbani,
 L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
 Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
 P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
 B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
 M., Shinada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
 R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L. G., Wyszewski, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

JOURNAL
MEDLINE
PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,

S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,

Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,

H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,

Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,

Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,

Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct

Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Takashi Ishikawa (Department of Surgery

2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama

236-0004 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

[illegible]

RESULT 15	BQ896389	linear	EST 16-AUG-2002
LOCUS	AGENCOURT_8743817	NIH_MGC_129	Mus musculus cDNA clone IMAGE:6389748
DEFINITION	5', mRNA sequence.		
ACCESSION	BQ896389		
VERSION	BQ896389.1	GI:22288403	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 942)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-r@mail.nih.gov		
	Tissue Procurement: Susan L. Sullivan, PhD.		
	cDNA Library Preparation: ResGen, Invitrogen Corp		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		

```

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13875 row: e column: 13
High quality sequence stop: 697.
Location/Qualifiers
1. .942
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6389748"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_129"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.ccdB; Site_1: EcoRV; Site_2: NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this
is a NIH MGC Library."
BASE COUNT 170 a 294 c 259 g 219 t

```

[illegible]

Search completed: December 20, 2003, 23:26:10
Job time : 1756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2003, 22:13:05 ; Search time 59 Seconds
(without alignments)
4099.626 Million cell updates/sec

Title: US-09-744-679-7

Perfect score: 548

Sequence: 1 cgcgagggtcgttgacc.....ccatcgttggttcacctg 548

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	548	100.0	3472	4	US-09-016-434-1328
2	548	100.0	3480	1	US-07-657-769B-68
3	548	100.0	3480	1	US-07-789-184-219
4	548	100.0	3480	1	US-08-475-263-219
5	548	100.0	3480	1	US-08-485-886-219
6	548	100.0	3480	2	US-08-477-362-219
7	548	100.0	3480	2	US-08-477-134-219
8	548	100.0	3480	3	US-08-473-489A-219
9	548	100.0	3480	3	US-08-485-695-219
10	548	100.0	3480	3	US-08-018-760-219
11	282.4	51.5	1764	1	US-08-313-553-12
12	282.4	51.5	1764	3	US-08-767-993-12
13	125.4	22.9	1300	4	US-08-795-876-31
14	125.4	22.9	1300	4	US-08-795-876-36
15	123.8	22.6	1300	4	US-08-795-876-37
16	123.6	22.6	1298	4	US-08-795-876-32
17	56	10.2	56	4	US-08-795-876-13
18	55	10.0	55	4	US-08-795-876-14
19	48.8	8.9	47981	4	US-09-679-279-1
20	47.6	8.7	54	4	US-08-795-876-12
21	44.6	8.1	192	1	US-08-644-664B-33
22	44.6	8.1	192	2	US-08-761-277A-33
23	40.8	7.4	72	1	US-08-093-741-50
24	40.8	7.4	72	1	US-08-720-012-50
25	40.6	7.4	189	1	US-08-644-664B-31
26	40.6	7.4	189	2	US-08-761-277A-31
27	39	7.1	72	1	US-08-093-741-49

c	28	39	7.1	72	1	US-08-720-012-49	Sequence 49, Appl
	29	38.6	7.0	8051	2	US-08-576-626A-2	Sequence 2, Appl
	30	38.2	7.0	1765	1	US-08-459-586-15	Sequence 15, Appl
	31	38.2	7.0	1765	2	US-08-282-696-15	Sequence 15, Appl
	32	38.2	7.0	18994	1	US-08-459-586-4	Sequence 4, Appl
	33	38.2	7.0	18994	2	US-08-282-696-4	Sequence 4, Appl
	34	38.2	7.0	35060	3	US-08-814-095-7	Sequence 7, Appl
	35	38	6.9	402	4	US-09-252-991A-12298	Sequence 12298, A
c	36	38	6.9	1302	4	US-09-252-991A-12124	Sequence 12124, A
c	37	38	6.9	2160	4	US-09-252-991A-12197	Sequence 12197, A
	38	38	6.9	2772	4	US-09-252-991A-12393	Sequence 12393, A
	39	38	6.9	3105	4	US-09-252-991A-12350	Sequence 12350, A
	40	37.8	6.9	30001	1	US-08-125-468-1	Sequence 1, Appl
	41	37.8	6.9	30001	2	US-08-474-933-1	Sequence 1, Appl
c	42	37.6	6.9	3177	1	US-08-042-747A-4	Sequence 4, Appl
	43	37.2	6.8	1910	4	US-09-149-476-297	Sequence 297, App
	44	37.2	6.8	1945	4	US-09-149-476-164	Sequence 164, App
	45	37	6.8	2065	3	US-08-335-865J-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-016-434-1328
; Sequence 1328, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1328:
SEQUENCE CHARACTERISTICS:
LENGTH: 3472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g339676
US-09-016-434-1328

Query Match 100.0%; Score 548; DB 4; Length 3472;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;

```
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 60
Db 64 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 123
QY 61 AAGACCGGCTCCCGGAGCAGCGCCGCGCAGAGGTGAGAGGAGAGGAGAGGAGAGGAGAGG 120
Db 124 AAGACCGGCTCCCGGAGCAGCGCCGCGCAGAGGTGAGAGGAGAGGAGAGGAGAGGAGAGG 183
QY 121 GGGGACGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGAGCAATGGGGCCGCGCGCTGC 180
Db 184 GGGGACGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGAGCAATGGGGCCGCGCGCTGC 243
QY 181 TGCTGGTGGCGGCTGCTTCAAGTCTGTGCGGCGCGCTGTGTCTGCCCGCAGCCCGGCGCC 240
Db 244 TGCTGGTGGCGGCTGCTTCAAGTCTGTGCGGCGCGCTGTGTCTGCCCGCAGCCCGGCGCC 303
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGACCCTTAGATCCCGCGTCAATTTCTTCTCAGGA 300
Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGACCCTTAGATCCCGCGTCAATTTCTTCTCAGGA 363
QY 301 ACCCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAGAGAGAGAGAGAGAGAGAG 360
Db 364 ACCCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAGAGAGAGAGAGAGAGAGAG 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAACAACTTCTTCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAACAACTTCTTCTGCAT 483
QY 421 TCATCTCAGAAAGATGCCCTCCGGATATTGACCAGCTCTCGGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAAAGATGCCCTCCGGATATTGACCAGCTCTCGGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611
```

RESULT 2

```
US-07-657-769B-68
; Sequence 68, Application US/07657769B
; Patent No. 5256766
;
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; TITLE OF INVENTION: RECOMBINANT RECEPTOR AND RELATED
; TITLE OF INVENTION: PHARMACEUTICALS
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IRELL & MANELLA
; STREET: 545 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/657,769B
; FILING DATE: 19910219
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2000-0502.00
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1503
US-07-657-769B-68
```

```
Query Match 100.0%; Score 548; DB 1; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 60
Db 64 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCCTGCCGCG 123
QY 61 AAGACCGGCTCCCGGAGCAGCGCCGCGCAGAGGTGAGAGGAGAGGAGAGGAGAGGAGAGG 120
Db 124 AAGACCGGCTCCCGGAGCAGCGCCGCGCAGAGGTGAGAGGAGAGGAGAGGAGAGGAGAGG 183
QY 121 GGGGACGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGAGCAATGGGGCCGCGCGCTGC 180
Db 184 GGGGACGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGAGCAATGGGGCCGCGCGCTGC 243
QY 181 TGCTGGTGGCGGCTGCTTCAAGTCTGTGCGGCGCGCTGTGTCTGCCCGCAGCCCGGCGCC 240
Db 244 TGCTGGTGGCGGCTGCTTCAAGTCTGTGCGGCGCGCTGTGTCTGCCCGCAGCCCGGCGCC 303
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCGCGTCAATTTCTTCTCAGGA 300
Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGCCACTTAGATCCCGCGTCAATTTCTTCTCAGGA 363
QY 301 ACCCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAGAGAGAGAGAGAGAGAGAG 360
Db 364 ACCCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAGAGAGAGAGAGAGAGAGAG 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAACAACTTCTTCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTCAAAACAACTTCTTCTGCAT 483
QY 421 TCATCTCAGAAAGATGCCCTCCGGATATTGACCAGCTCTCGGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAAAGATGCCCTCCGGATATTGACCAGCTCTCGGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGGAGTGTGTTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611
```

RESULT 3

```
US-07-789-184-219
; Sequence 219, Application US/07789184
; Patent No. 5688768
;
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
```


; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,184
; FILING DATE: 19911107
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-07-789-184-219

Query Match 100.0%; Score 548; DB 1; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCGTGGGACCCCTGGCTCTGCCTGCCGCG 60
Db 64 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCGTGGGACCCCTGGCTCTGCCTGCCGCG 123
QY 61 AAGACCGGCTCCCGACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 124 AAGACCGGCTCCCGACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183
QY 121 GGGGAGCGCTCCCGAGCAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 184 GGGGAGCGCTCCCGAGCAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 243
QY 181 TGCTGGTGGCGGCTGCTTTCAGTCTGTGCGGCGGCTGTTGTCTGCGGCGGAGGAGGAGGAG 240
Db 244 TGCTGGTGGCGGCTGCTTTCAGTCTGTGCGGCGGCTGTTGTCTGCGGCGGAGGAGGAGGAG 303
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGACCTTACCTTACCTTACCTTACCTTACCTTAC 300
Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGACCTTACCTTACCTTACCTTACCTTACCTTAC 363
QY 301 ACCCAATGATAAATATGAACCATTTTGGAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 364 ACCCAATGATAAATATGAACCATTTTGGAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGG 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACTTCTTCCCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACTTCTTCCCAT 483
QY 421 TCATCTCAGAGATGCTCCCGGATATTGACAGCTCCTGGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAGATGCTCCCGGATATTGACAGCTCCTGGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGAGTGTGTTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGAGTGTGTTGTAGTCAGCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548

Db 604 TCATCCTG 611
RESULT 4
US-08-475-263-219
; Sequence 219, Application US/08475263
; Patent No. 5759994
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,263
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-475-263-219

Query Match 100.0%; Score 548; DB 1; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCGTGGGACCCCTGGCTCTGCCTGCCGCG 60
Db 64 CGCCGAGGGTCGCTTGGACCCCTGATCTTACCGTGGGACCCCTGGCTCTGCCTGCCGCG 123
QY 61 AAGACCGGCTCCCGACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 124 AAGACCGGCTCCCGACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183
QY 121 GGGGAGCGCTCCCGAGCAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 184 GGGGAGCGCTCCCGAGCAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 243
QY 181 TGCTGGTGGCGGCTGCTTTCAGTCTGTGCGGCGGCTGTTGTCTGCGGCGGAGGAGGAGGAG 240
Db 244 TGCTGGTGGCGGCTGCTTTCAGTCTGTGCGGCGGCTGTTGTCTGCGGCGGAGGAGGAGGAG 303
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGACCTTACCTTACCTTACCTTACCTTACCTTAC 300
Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGACCTTACCTTACCTTACCTTACCTTACCTTAC 363

QY 301 ACCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAAAAATGAAAAGTGGGTAA 360
Db 364 ACCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAAAAATGAAAAGTGGGTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAACITTCCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAACITTCCTGCAT 483
QY 421 TCATCTCAGAAAGATGCTCCGATATTTGACAGCTCCCTGCTGACACTCTTTGTCCTCAT 480
Db 484 TCATCTCAGAAAGATGCTCCGATATTTGACAGCTCCCTGCTGACACTCTTTGTCCTCAT 543
QY 481 CTGTGTACACGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 5
US-08-485-886-219
; Sequence 219, Application US/08485886
; Patent No. 5798248
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499

US-08-485-886-219
Query Match 100.0%; Score 548; DB 1; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCGCTGCCGCG 60
Db 64 CGCCGAGGTGCTTGGACCCCTGATCTTACCCGTGGGCACCCCTGCGCTCTGCGCTGCCGCG 123
QY 61 AAGACCGGCTCCCCGACCCCGCAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCCGAGGC 120
Db 124 AAGACCGGCTCCCCGACCCCGCAGAGTCAAGAGAGAGGGTGAAGCGGAGCAGCCCCGAGGC 183
QY 121 GGGGAGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGGAACAATGGGGCCCGCGGCGTGC 180
Db 184 GGGGAGCTCCCGGAGCAGCGCCGCGCAGAGCCCGGGAACAATGGGGCCCGCGGCGTGC 243
QY 181 TGCTGTGGCCGCTGCTTCAAGTGTGCGGCCCGCTGTGTCTGCCCGCAGCCCCGGGCC 240
Db 244 TGCTGTGGCCGCTGCTTCAAGTGTGCGGCCCGCTGTGTCTGCCCGCAGCCCCGGGCC 303
QY 241 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGA 300
Db 304 GCAGGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCCGGTCATTTCTTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAAAAATGAAAAGTGGGTAA 360
Db 364 ACCCAATGATAAATATGAACCAATTTGGGAGGATGAGGAGAAAAATGAAAAGTGGGTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAACITTCCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAAACAACITTCCTGCAT 483
QY 421 TCATCTCAGAAAGATGCTCCGATATTTGACAGCTCCCTGCTGACACTCTTTGTCCTCAT 480
Db 484 TCATCTCAGAAAGATGCTCCGATATTTGACAGCTCCCTGCTGACACTCTTTGTCCTCAT 543
QY 481 CTGTGTACACGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACGGAGTGTGTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 6
US-08-477-362-219
; Sequence 219, Application US/08477362
; Patent No. 5849507
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,362
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959

Db 544 CTGTGTACACCGAGTGTGTTAGTACGCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 8
US-08-473-489A-219
; Sequence 219, Application US/08473489A
; Patent No. 6024936
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,489A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,184
; FILING DATE: 1991-11-07
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
US-08-473-489A-219

Query Match 100.0%; Score 548; DB 3; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGAGGTCGCTTGACCTGATCTTACCCGTGGGACCCCTGCGCTCTGCTGCCGCG 60
Db 64 CGCCGAGGTCGCTTGACCTGATCTTACCCGTGGGACCCCTGCGCTCTGCTGCCGCG 123
QY 61 AAGACCGCTCCCGACCCGAGAGTCAAGGAGAGGTTGAAGCGGAGAGCCCGAGGC 120
Db 124 AAGACCGCTCCCGACCCGAGAGTCAAGGAGAGGTTGAAGCGGAGAGCCCGAGGC 183
QY 121 GGGGAGCTCCCGAGAGCGCCGCGAGAGCCCGGAGCAATGGGGCGCGCGGCTGC 180
Db 184 GGGGAGCTCCCGAGAGCGCCGCGAGAGCCCGGAGCAATGGGGCGCGCGGCTGC 243
QY 181 TGCTGTGGCGCGCTGCTTACGTTGTGCGGCGCGCTGTGTCTGCGCGACCCGCGCC 240

Db 244 TGCTGTGGCGCGCTGCTTCACTCTGTGCGCGCGCTGTGTGCTGCGCGCGCGCC 303
QY 241 GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCATTTCTTCTCAGGA 300
Db 304 GCAGGCCAGAAATCAAAAGCAACAAATGCCACCTTAGATCCCGGTCATTTCTTCTCAGGA 363
QY 301 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAATGAAAGTGGTTAA 360
Db 364 ACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATATAAGCAGTCTCTTCAAAAACAACTTCTCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATATAAGCAGTCTCTTCAAAAACAACTTCTCTGCAT 483
QY 421 TCATCTCAGAAAGATGCTCCGGATATTGACACGCTCCTGGCTGACACTCTTTGTCCCAT 480
Db 484 TCATCTCAGAAAGATGCTCCGGATATTGACACGCTCCTGGCTGACACTCTTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGTGTGTTAGTACGCTCCCACTAAACATCATGGCCATCGTTGTGT 540
Db 544 CTGTGTACACCGGAGTGTGTTAGTACGCTCCCACTAAACATCATGGCCATCGTTGTGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 9
US-08-485-695-219
; Sequence 219, Application US/08485695
; Patent No. 6124101
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,695
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

; LOCATION: 225..1499
US-08-485-695-219
Query Match 100.0%; Score 548; DB 3; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGTCGCTTGGACCTGATCTTACCCGTGGGACCCCTGCGCTCTGCGTCCGCG 60
Db 64 CGCCGAGGTCGCTTGGACCTGATCTTACCCGTGGGACCCCTGCGCTCTGCGTCCGCG 123
QY 61 AAGACCGGCTCCCGACCCGAGAGTCAAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
Db 124 AAGACCGGCTCCCGACCCGAGAGTCAAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 183
QY 121 GGGGAGCCTCCCGAGCAGCGCGCGCAGAGCCCGGACAATGGGGCCGCGCGGTGC 180
Db 184 GGGGAGCCTCCCGAGCAGCGCGCGCAGAGCCCGGACAATGGGGCCGCGCGGTGC 243
QY 181 TGCTGTGGCGCGCTTTCAGTCTGTGCGGCGCGCTGTGTCTGCCGACCCCGGGCCC 240
Db 244 TGCTGTGGCGCGCTTTCAGTCTGTGCGGCGCGCTGTGTCTGCCGACCCCGGGCCC 303
QY 241 GCAGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTCTCAGGA 300
Db 304 GCAGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAATGAAAGTGGTTAA 360
Db 364 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 483
QY 421 TCATCTCAGAAAGATGCCTCCGATATTTGACAGCTCTCCAGCTGACACTCTTGTCCCAT 480
Db 484 TCATCTCAGAAAGATGCCTCCGATATTTGACAGCTCTCCAGCTGACACTCTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGT 540
Db 544 CTGTGTACACCGGAGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 10
US-08-018-760-219
; Sequence 219, Application US/08018760
; Patent No. 6197541
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/018,760
; FILING DATE: 17-FEB-1993

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
US-08-018-760-219

Query Match 100.0%; Score 548; DB 3; Length 3480;
Best Local Similarity 100.0%; Pred. No. 2.3e-134;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGCTCTGCGTCCGCG 60
Db 64 CGCCGAGGTCGCTTGGACCCCTGATCTTACCCGTGGGACCCCTGCGCTCTGCGTCCGCG 123
QY 61 AAGACCGGCTCCCGACCCGAGAGTCAAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 120
Db 124 AAGACCGGCTCCCGACCCGAGAGTCAAGGAGAGGGTGAAGCGGAGCAGCCCGAGGC 183
QY 121 GGGGAGCCTCCCGAGCAGCGCGCGCAGAGCCCGGACAATGGGGCCGCGCGGTGC 180
Db 184 GGGGAGCCTCCCGAGCAGCGCGCGCAGAGCCCGGACAATGGGGCCGCGCGGTGC 243
QY 181 TGCTGTGGCGCGCTTTCAGTCTGTGCGGCGCGCTGTGTCTGCCGACCCCGGGCCC 240
Db 244 TGCTGTGGCGCGCTTTCAGTCTGTGCGGCGCGCTGTGTCTGCCGACCCCGGGCCC 303
QY 241 GCAGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTCTCAGGA 300
Db 304 GCAGCCAGAAATCAAAAGCAAAATGCCACCTTAGATCCCGGTCAATTTCTCTCAGGA 363
QY 301 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAATGAAAGTGGTTAA 360
Db 364 ACCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAAATGAAAGTGGTTAA 423
QY 361 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 420
Db 424 CTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAAACAACCTTCTGCAT 483
QY 421 TCATCTCAGAAAGATGCCTCCGATATTTGACAGCTCTCCAGCTGACACTCTTGTCCCAT 480
Db 484 TCATCTCAGAAAGATGCCTCCGATATTTGACAGCTCTCCAGCTGACACTCTTGTCCCAT 543
QY 481 CTGTGTACACCGGAGT 540
Db 544 CTGTGTACACCGGAGT 603
QY 541 TCATCCTG 548
Db 604 TCATCCTG 611

RESULT 11
US-08-313-553-12
; Sequence 12, Application US/08313553
; Patent No. 5641650
; GENERAL INFORMATION:


```
; APPLICANT: TURNER, George J.
; APPLICANT: BETLACH, Mary C.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
; TITLE OF INVENTION: IN HALOBACTERIA
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,553
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,662
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-57669/WHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 436..462
; OTHER INFORMATION: /note= "Sequence encoding
; Patent No. 5641650
; OTHER INFORMATION: polyaspartic acid."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 463..465
; OTHER INFORMATION: /note= "Codon encoding the
; OTHER INFORMATION: N-terminal amino acid of the human thrombin
; OTHER INFORMATION: receptor protein."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1630..1632
; OTHER INFORMATION: /note= "Codon encoding the
; OTHER INFORMATION: C-terminal amino acid of the human thrombin
; OTHER INFORMATION: receptor protein."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 1633..1650
; OTHER INFORMATION: /note= "Sequence encoding
; Patent No. 5641650
; OTHER INFORMATION: polyhistidine."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 648..656
; OTHER INFORMATION: /note= "Deleted AlwNI restriction
; OTHER INFORMATION: site."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 893..898
; OTHER INFORMATION: /note= "Deleted PstI restriction
; OTHER INFORMATION: site."
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1301..1309
; OTHER INFORMATION: /note= "Deleted AlwNI restriction
; OTHER INFORMATION: site."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1394..1402
; OTHER INFORMATION: /note= "Deleted AlwNI restriction
; OTHER INFORMATION: site."
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 374
; OTHER INFORMATION: /note= "RNA start site."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(1671, "")
; OTHER INFORMATION: /note= "C to T mutation removes
; OTHER INFORMATION: AlwNI site."
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 376..1650
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 376..414
; OTHER INFORMATION: /note= "Bacteriorhodopsin
; OTHER INFORMATION: pre-sequence."
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 1672..1674
; OTHER INFORMATION: /note= "Bacteriorhodopsin stop
; OTHER INFORMATION: codon."
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 1651..1653
; OTHER INFORMATION: /note= "Thrombin stop codon."
; US-08-313-553-12
;
Query Match 51.5%; Score 282.4; DB 1; Length 1764;
Best Local Similarity 99.6%; Pred. No. 8.1e-65;
Matches 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
Qy 265 ATGCCACCTTAGATCCCGGTCAATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAT 324
Db 461 AGCCACCTTAGATCCCGGTCAATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAT 520
;
Qy 325 TTGGGAGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCA 384
Db 521 TTGGGAGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGATTAGTCTCCATCA 580
;
Qy 385 ATAAAAGCAGTCCTTCAAAAAACAACCTTCTCGCATTCATCTCAGAAGATGCTCCGGAT 444
Db 581 ATAAAAGCAGTCCTTCAAAAAACAACCTTCTCGCATTCATCTCAGAAGATGCTCCGGAT 640
;
Qy 445 ATTTGACCAGTCCTCGGTGACACTCTTTTGTCCTCATCTGTGTACACCGGAGTGTGTAG 504
Db 641 ATTTGACCAGTCCTCGGTGACACTCTTTTGTCCTCATCTGTGTACACCGGAGTGTGTAG 700
;
Qy 505 TCAGCCTCCCACTAAACATCATGGCCATCGTTGTGTTTCATCCTG 548
Db 701 TCAGCCTCCCACTAAACATCATGGCCATCGTTGTGTTTCATCCTG 744
;
RESULT 12
US-08-767-993-12
; Sequence 12, Application US/08767993
; Patent No. 6010885
; GENERAL INFORMATION:
; APPLICANT: TURNER, George J.
; APPLICANT: BETLACH, Mary C.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
; TITLE OF INVENTION: IN HALOBACTERIA
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,876
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BRAMAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1280
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-795-876-31

Query Match 22.9%; Score 125.4; DB 4; Length 1300;
Best Local Similarity 86.8%; Pred. No. 1.1e-23;
Matches 138; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 314 ATATGAACCACTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAATACAGATT 373
Db ||||| |
155 ATATGCCACCTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAATACAGATT 214
QY 374 AGTCTCCATCAATAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA 433
Db ||||| |
215 AGTCTCCATCAATAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA 274
QY 434 TGCCTCCGGATATTGACCAGCTCCTCGGCTGACACTCTT 472
Db ||||| |
275 TGCCTCCGGTACCTGGGGTACAACATCCTCAGAGTCCT 313

RESULT 14
US-08-795-876-36
Sequence 36, Application US/08795876
Patent No. 6403305
GENERAL INFORMATION:
APPLICANT: Gershengorn, Marvin C.
APPLICANT: Geras-Raaka, Elizabeth
APPLICANT: Nussenzweig, Daniel R.
TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN
TITLE OF INVENTION: COUPLED RECEPTORS
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,876
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BRAMAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1280
TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-795-876-36

Query Match 22.9%; Score 125.4; DB 4; Length 1300;
Best Local Similarity 86.8%; Pred. No. 1.1e-23;
Matches 138; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 314 ATATGAACCACTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAATACAGATT 373
Db ||||| |
155 ATATGCCACCTTTGGGAGGATGAGGAGAAAAATGAAAGTGGTTAACTGAATACAGATT 214
QY 374 AGTCTCCATCAATAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA 433
Db ||||| |
215 AGTCTCCATCAATAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA 274
QY 434 TGCCTCCGGATATTGACCAGCTCCTCGGCTGACACTCTT 472
Db ||||| |
275 TGCCTCCGGTACCTGGGGTACAACATCCTCAGAGTCCT 313

RESULT 15
US-08-795-876-37/c
Sequence 37, Application US/08795876
Patent No. 6403305
GENERAL INFORMATION:
APPLICANT: Gershengorn, Marvin C.
APPLICANT: Geras-Raaka, Elizabeth
APPLICANT: Nussenzweig, Daniel R.
TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN
TITLE OF INVENTION: COUPLED RECEPTORS
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,876
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BRAMAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1280
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-795-876-37

Query Match 22.6%; Score 123.8; DB 4; Length 1300;
Best Local Similarity 86.2%; Pred. No. 2.9e-23;
Matches 137; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY	314	ATATGAACCATTTTGGGAGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGATT	373
Db	1150	ATATGCCACCTTTTGGGAGGATGAGGAGAAAAAATGAAAGTGGTTAACTGAATACAGTTT	1091
QY	374	AGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA	433
Db	1090	AGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAAACAACCTTCCTGCATTTCATCTCAGAAGA	1031
QY	434	TGCCTCCGGATATTTGACCCAGCTCCTGGGCTGACACTCTT	472
Db	1030	TGCCTCCGGTTACCTGGGGTACAACATCCTCAGAGTCCT	992

Search completed: December 20, 2003, 23:32:58
Job time : 61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2003, 21:56:50 ; Search time 258 Seconds
(without alignments)
7077.703 Million cell updates/sec

Title: US-09-744-679-7

Perfect score: 548

Sequence: 1 cgccgagggtcgcttgacc.....ccatgtgtgttcctcgtg 548

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	548	100.0	3592	15	US-10-225-567A-323
2	548	100.0	3592	15	US-10-177-293-129
3	445	81.2	3299	15	US-10-208-408-24
4	387	70.6	1278	13	US-10-176-464A-2
5	301.6	55.0	6203	13	US-10-176-464A-1
6	236.2	43.1	405	11	US-09-918-995-16689
7	148.2	27.0	6436	13	US-10-311-455-654
8	146.6	26.8	6436	13	US-10-311-455-653
9	102.4	18.7	1440	13	US-10-176-464A-66
10	44.6	8.1	192	10	US-09-925-664-33
11	44.4	8.1	897	13	US-10-029-386-24941
12	43.2	7.9	309	15	US-10-156-761-610
13	43.2	7.9	9025608	15	US-10-156-761-1
14	42.4	7.7	846	15	US-10-156-761-7276
15	42.4	7.7	9025608	15	US-10-156-761-1

16	42	7.7	1737	13	US-10-029-386-25063	Sequence 25063, A
17	40.8	7.4	729	13	US-10-027-632-126950	Sequence 126950, A
18	40.8	7.4	729	14	US-10-027-632-126950	Sequence 126950, A
19	40.8	7.4	737	13	US-10-027-632-16043	Sequence 16043, A
20	40.8	7.4	737	14	US-10-027-632-16043	Sequence 16043, A
21	40.8	7.4	1323	15	US-10-156-761-5360	Sequence 5360, Ap
22	40.6	7.4	189	10	US-09-925-664-31	Sequence 31, Appl
23	40.6	7.4	9369	13	US-10-200-562-190	Sequence 190, App
24	40.6	7.4	9369	13	US-10-237-551-190	Sequence 190, App
25	40.6	7.4	9369	13	US-10-237-551-247	Sequence 247, App
26	40.6	7.4	154746	13	US-09-827-688-8	Sequence 8, Appli
27	40.4	7.4	1738	13	US-10-029-386-25470	Sequence 25470, A
28	40.2	7.3	592	13	US-10-027-632-126949	Sequence 126949, A
29	40.2	7.3	592	14	US-10-027-632-126949	Sequence 126949, A
30	40.2	7.3	656	13	US-10-027-632-105398	Sequence 105398, A
31	40.2	7.3	656	14	US-10-027-632-105398	Sequence 105398, A
32	40.2	7.3	32195	9	US-09-764-870-611	Sequence 611, App
33	40.2	7.3	32195	9	US-09-764-870-617	Sequence 617, App
34	40.2	7.3	32195	9	US-09-764-869-1605	Sequence 1605, Ap
35	40.2	7.3	32195	15	US-10-125-540-611	Sequence 611, App
36	40.2	7.3	32195	15	US-10-125-540-617	Sequence 617, App
37	40.2	7.3	32195	15	US-10-091-504-1605	Sequence 1605, Ap
38	39.8	7.3	2793	15	US-10-156-761-2889	Sequence 2889, Ap
39	39.8	7.3	125746	15	US-10-156-761-15102	Sequence 15102, A
40	39.6	7.2	409	10	US-09-960-352-13321	Sequence 13321, A
41	39.6	7.2	594	13	US-10-140-472-10	Sequence 10, Appl
42	39.6	7.2	594	13	US-10-141-761-10	Sequence 10, Appl
43	39.6	7.2	594	13	US-10-142-885-10	Sequence 10, Appl
44	39.6	7.2	594	13	US-10-158-790-10	Sequence 10, Appl
45	39.6	7.2	594	13	US-10-137-871-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-225-567A-323
; Sequence 323, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 323
; LENGTH: 3592
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-225-567A-323

Query Match	100.0%	Score 548;	DB 15;	Length 3592;
Best Local Similarity	100.0%	Pred. No. 7.7e-158;		
Matches 548;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGCCGAGGTCGCTTGACCCCTGATCTTACCGTGGCACCCTGCGCTTGCCTGCCGCG 60		
DB	184	CGCCGAGGTCGCTTGACCCCTGATCTTACCGTGGCACCCTGCGCTTGCCTGCCGCG 243		
QY	61	AAGACCGCTCCCGACCCGACAGAGTCAAGTCAAGAGAGAGGTTGAAGCGGAGAGCCGAGGC 120		
DB	244	AAGACCGCTCCCGACCCGACAGAGTCAAGTCAAGAGAGAGGTTGAAGCGGAGAGCCGAGGC 303		
QY	121	GGGGAGCCTCCCGAGAGAGCCCGCGAGAGCCCGGACCAATGGGCGCGCGGCTGC 180		
DB	304	GGGGAGCCTCCCGAGAGAGCCCGCGAGAGCCCGGACCAATGGGCGCGCGGCTGC 363		


```
Query Match      81.2%; Score 445; DB 15; Length 3299;
Best Local Similarity 100.0%; Pred. No. 3.8e-126;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 GGGGAGCAGCCCGGAGCGGGGAGCCCTCCCGGAGCAGCGCGGAGAGCCCGGAGCAAT 163
Db 1 GGGGAGCAGCCCGGAGCGGGGAGCCCTCCCGGAGCAGCGCGGAGAGCCCGGAGCAAT 60

QY 164 GGGGCGCGGGGCTGCTGCTGGTGGCGCGCTGCTTCACTGCTGCTGGCGCGCTGCTGTC 223
Db 61 GGGGCGCGGGGCTGCTGCTGGTGGCGCGCTGCTTCACTGCTGCTGGCGCGCTGCTGTC 120

QY 224 TGCCCGCACCCCGGCGCGGAGCCAGGATCAAAAGCAACAATGCCACCTTAGATCCCGG 283
Db 121 TGCCCGCACCCCGGCGCGGAGCCAGGATCAAAAGCAACAATGCCACCTTAGATCCCGG 180

QY 284 GTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAA 343
Db 181 GTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAGAA 240

QY 344 AAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGAGCATCTCTTCA 403
Db 241 AAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGAGCATCTCTTCA 300

QY 404 AAAACAACTTCTGCTGCTCATCTCAGAAAGATGCTCCGGATATTTGACCAAGCTCCGGCT 463
Db 301 AAAACAACTTCTGCTGCTCATCTCAGAAAGATGCTCCGGATATTTGACCAAGCTCCGGCT 360

QY 464 GACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACAT 523
Db 361 GACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAACAT 420

QY 524 CATGGCCATCGTTGTGTTTCATCCTG 548
Db 421 CATGGCCATCGTTGTGTTTCATCCTG 445

RESULT 4
US-10-176-464A-2
; Sequence 2, Application US/10176464A
; Publication No. US20030165902A1
; GENERAL INFORMATION:
; APPLICANT: Bieglecki, Karyn
; APPLICANT: Lee, Helen
; APPLICANT: Messer, Chad
; APPLICANT: Monroe, Glen
; TITLE OF INVENTION: HAPLOTYPES OF THE F2R GENE
; FILE REFERENCE: F2R MMH-1457US
; CURRENT APPLICATION NUMBER: US/10/176,464A
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/30876
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/236,603
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-464A-2

Query Match      70.6%; Score 387; DB 13; Length 1278;
Best Local Similarity 100.0%; Pred. No. 1.7e-108;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ATGGGGCCGCGGCGGCTGCTGCTGGTGGCGCGCTGCTTCACTGCTGCGGCGCGCTGTTG 221
Db 1 ATGGGGCCGCGGCGGCTGCTGCTGGTGGCGCGCTGCTTCACTGCTGCGGCGCGCTGTTG 60

QY 222 TCTGCCCGCACCCGCGGCGCGGAGGATCAAAAGCAACAATGCCACCTTAGATCC 281
Db 61 TCTGCCCGCACCCGCGGCGCGGAGGATCAAAAGCAACAATGCCACCTTAGATCC 120
```

```
QY 282 CGGTCAATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAG 341
Db 121 CGGTCAATTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAGGATGAGGAG 180

QY 342 AAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGAGCATCTCTT 401
Db 181 AAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGAGCATCTCTT 240

QY 402 CAAAAACAACCTTCTGCTGCTCATCTCAGAAAGATGCTCCGGATATTTGACCAAGCTCCTGG 461
Db 241 CAAAAACAACCTTCTGCTGCTCATCTCAGAAAGATGCTCCGGATATTTGACCAAGCTCCTGG 300

QY 462 CTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAAC 521
Db 301 CTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTTAGTACAGCTCCCACTAAAC 360

QY 522 ATCATGGCCATCGTTGTGTTTCATCCTG 548
Db 361 ATCATGGCCATCGTTGTGTTTCATCCTG 387

RESULT 5
US-10-176-464A-1
; Sequence 1, Application US/10176464A
; Publication No. US20030165902A1
; GENERAL INFORMATION:
; APPLICANT: Bieglecki, Karyn
; APPLICANT: Lee, Helen
; APPLICANT: Messer, Chad
; APPLICANT: Monroe, Glen
; TITLE OF INVENTION: HAPLOTYPES OF THE F2R GENE
; FILE REFERENCE: F2R MMH-1457US
; CURRENT APPLICATION NUMBER: US/10/176,464A
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/30876
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/236,603
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6203
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1318)..(1318)
; OTHER INFORMATION: PS1: polymorphic base cytosine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1467)..(1467)
; OTHER INFORMATION: PS2: polymorphic base cytosine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1475)..(1475)
; OTHER INFORMATION: PS3: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1499)..(1499)
; OTHER INFORMATION: PS4: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1519)..(1519)
; OTHER INFORMATION: PS5: polymorphic base thymine or cytosine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1584)..(1584)
; OTHER INFORMATION: PS6: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2741)..(2840)
; OTHER INFORMATION: N's represent unknown nucleotides
```


Db 1069 TACTAATAACCGCTACTTCAATCTATACGACCGCTATTATCTACCCGACCCGACCC 1010
Qy 241 GCAGCCAGCAATCAAAAGCAACAAA 265
Db 1009 GCAACCAATAATAAAATACACGAA 985

RESULT 8
US-10-311-455-653
; Sequence 653, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Polymorphic Base Cytosine or Adenine
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 653
; LENGTH: 6436
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-653

Query Match 26.8%; Score 146.6; DB 13; Length 6436;
Best Local Similarity 74.3%; Pred. No. 4.1e-34;
Matches 185; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 1 CGCCGAGGTCGCTTGGACCTGATCTTACCCGTTGGCACCCTGCGCTCTGCTGCGCG 60
Db 5188 CGTCGAGGTCGCTTGGATTTTATTCGTTGCTATTTTTCGTTTGTTCGCG 5247
QY 61 AAGACCGGTCCTCCCGAGCAGCCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 5248 AAGATCGGTTTTCGATTCTGTAAGTTAGGAGAGAGGAGGAGGAGGAGGAGGAGG 5307
QY 121 GGGGAGGTCCTCCCGAGCAGCCGCGAGAGCCGCGAGACAAATGGGGCCGCGGCTGC 180
Db 5308 GGGGTAGTTTTCGAGTAGTCGCTGCGTAGAGTTCCGGATAATGGGGTCGCGGCTGT 5367
QY 181 TGCTGGTGGCCGCTGCTTCACTGTGTGCGGCGGCTGTGTGTGCTGCGGCGGCGG 240
Db 5368 TGTGGTGGTGGTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5427
QY 241 GCAGGCCAG 249
Db 5428 GTAGGTTAG 5436

RESULT 9
US-10-176-464A-66
; Sequence 66, Application US/10176464A
; Publication No. US20030165902A1
; GENERAL INFORMATION:
; APPLICANT: Bieglecki, Karyn
; APPLICANT: Lee, Helen
; APPLICANT: Messer, Chad
; APPLICANT: Monroe, Glen
; TITLE OF INVENTION: HAPLOTYPES OF THE F2R GENE
; FILE REFERENCE: F2R MWH-1457US
; CURRENT APPLICATION NUMBER: US/10/176,464A

; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/30876
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/236,603
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (30)..(30)
; OTHER INFORMATION: PS1: polymorphic base cytosine or adenine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (61)..(120)
; OTHER INFORMATION: N's represent nucleotides between PS1 and PS2
; FEATURE:
; NAME/KEY: allele
; LOCATION: (150)..(150)
; OTHER INFORMATION: PS2: polymorphic base cytosine or adenine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (181)..(240)
; OTHER INFORMATION: N's represent nucleotides between PS2 and PS3
; FEATURE:
; NAME/KEY: allele
; LOCATION: (270)..(270)
; OTHER INFORMATION: PS3: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (301)..(360)
; OTHER INFORMATION: N's represent nucleotides between PS3 and PS4
; FEATURE:
; NAME/KEY: allele
; LOCATION: (390)..(390)
; OTHER INFORMATION: PS4: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (421)..(480)
; OTHER INFORMATION: N's represent nucleotides between PS4 and PS5
; FEATURE:
; NAME/KEY: allele
; LOCATION: (510)..(510)
; OTHER INFORMATION: PS5: polymorphic base thymine or cytosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (541)..(600)
; OTHER INFORMATION: N's represent nucleotides between PS5 and PS6
; FEATURE:
; NAME/KEY: allele
; LOCATION: (630)..(630)
; OTHER INFORMATION: PS6: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (661)..(720)
; OTHER INFORMATION: N's represent nucleotides between PS6 and PS7
; FEATURE:
; NAME/KEY: allele
; LOCATION: (750)..(750)
; OTHER INFORMATION: PS7: polymorphic base adenine or guanine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (781)..(840)
; OTHER INFORMATION: N's represent nucleotides between PS7 and PS8
; FEATURE:
; NAME/KEY: allele
; LOCATION: (870)..(870)
; OTHER INFORMATION: PS8: polymorphic base guanine or adenine
; FEATURE:
; NAME/KEY: misc feature

